Capitolo III

APPLICAZIONI DEL CALCOLO DIFFERENZIALE

1. Funzioni omogenee.

Sia $A \subset R^n$ un insieme di raggi uscenti dall'origine cioè un insieme A di punti tale che se $\underline{x} \in A$, anche $t\underline{x} \in A$ per ogni $t > 0$.
Sia f una funzione a valori reali definita in A. Diremo che f è **positivamente omogenea** di grado α in A quando per ogni $\underline{x} \in A$ risulta

$$(1.1) \qquad f(t\underline{x}) = t^\alpha f(\underline{x}) \qquad \forall t > 0 .$$

Esempi

1) Il polinomio (omogeneo!) di secondo grado $f(x_1,x_2)=x_1^2+x_1x_2$ è positivamente omogeneo di grado 2 in R^2.

2) La funzione $f(x_1,x_2)=\frac{1}{x_1^2+x_2^2}$ è positivamente omogenea di grado -2 in $R^2-\{\underline{0}\}$.

3) La funzione $f(x_1,x_2,x_3)=\operatorname{sen}\frac{x_1+x_2}{x_3}$ è positivalente omogenea di grado 0 in R^3 privato della retta $x_3=0$.

4) La funzione $f(x_1)=\sqrt{x_1}$ è positivamente omogenea di grado $\frac{1}{2}$ sul semiasse $x_1 \geq 0$.

Osservazioni - 1) f è positivamente omogenea di grado α in A se e solo se $\frac{f(t\underline{x})}{t^\alpha}$ è indipendente da t in A.

2) Se $\alpha=0$, f è costante su ogni raggio uscente dall'origine. Se $\alpha\neq0$, al variare di $\underline{x}$ su un raggio uscente dall'origine, f ha segno costante e assume valori proporzionali a $\|\underline{x}\|^\alpha$.
(con costante di propor. $f(\frac{x}{\|x\|})$)

Esaminiamo ora le funzioni positivamente omogenee dal punto di vista della loro differenziabilità .

<u>Teorema</u> 1.1 - Se f è una funzione positivamente omogenea di grado α , differenziabile in un punto $\underline{x}^o \neq \underline{0}$, allora essa è differenziabile in ogni punto del raggio $(\underline{0},\underline{x}^o)$ e risulta

$$(1.2) \qquad \text{grad } f(t\underline{x}^o)=t^{\alpha-1} \text{ grad } f(\underline{x}^o) \quad \forall t>0 \ .$$

Infatti, per ogni $t>0$ si ha

$$\begin{aligned} f(t\underline{x}^o+\Delta\underline{x})-f(t\underline{x}^o) &= t^{\alpha}\left\{f\left(\underline{x}^o+\frac{\Delta\underline{x}}{t}\right)-f(\underline{x}^o)\right\} \\ &= t^{\alpha}\left\{\text{grad } f(\underline{x}^o)\ \frac{\Delta\underline{x}}{t} +o(\|\Delta\underline{x}\|)\right\} \\ &= t^{\alpha-1}\text{grad } f(\underline{x}^o)\ \Delta\underline{x}+o(\|\Delta\underline{x}\|) \end{aligned}$$

e cioè l'asserto.

Dalla (1.2)si ricava immediatamente il

<u>Corollario</u> 1.1 - Le derivate prime di una funzione f positivamente omogena di grado α e differenziabile sono positivamente omogenee di grado α -1.

Più in generale vale il

<u>Teorema</u> 1.2 - Se f è positivamente omogenea di grado α e derivabile secondo il versore $\underline{v}$ in un punto $\underline{x}^o$, allora esiste la derivata secondo il versore $\underline{v}$ in ogni punto del raggio $(\underline{0},\underline{x}^o)$ e tale derivata è positivamente omogenea di grado α-1 .

Infatti per ogni $t>0$ si ha

$$\frac{f(t\underline{x}^o+\tau\underline{v})-f(t\underline{x}^o)}{\tau} = t^{\alpha}\,\frac{f(\underline{x}^o+\frac{\tau}{t}\,\underline{v})-f(\underline{x}^o)}{\tau} = t^{\alpha-1}\,\frac{f(\underline{x}^o+\frac{\tau}{t}\,\underline{v})-f(\underline{x}^o)}{\frac{\tau}{t}}$$

da cui,facendo tendere τ a zero, segue l'asserto.

<u>Teorema</u> 1.3 (di <u>Eulero</u>) - La funzione $f:R^n \to R$ sia differenziabile in un aperto A, formato da raggi uscenti dall'origine. Allora f è positivamente omogenea di grado α in A se e solo se risulta

$$(1.3) \qquad \text{grad } f(\underline{x})\cdot\underline{x} = \alpha f(\underline{x}) \qquad \forall\, \underline{x}\in A \ .$$

Infatti, osserviamo preventivamente che, qualunque sia $\underline{x}\in A$, la funzione $\varphi(t)=\frac{f(t\underline{x})}{t^{\alpha}}$ è derivabile rispetto a t per ogni $t>0$ e risulta

$$(1.4) \qquad \varphi'(t) = \frac{t^{\alpha}\text{grad } f(t\underline{x})\cdot\underline{x}-\alpha t^{\alpha-1} f(t\underline{x})}{t^{2\alpha}} = \frac{\text{grad } f(t\underline{x})\cdot t\underline{x}-\alpha f(t\underline{x})}{t^{\alpha+1}} \ .$$

Allora, se vale (1.3),risulta $\varphi'(t)=0$; quindi $\varphi(t)$ è costante e f è positivamente omogenea di grado α per l'osservazione 1).

Viceversa, se f è positivamente omogenea di grado α, sempre per l'osservazione 1), $\varphi(t)$ è costante e quindi $\varphi'(t)=0$ per ogni t . Da (1.4) si ottiene allora

$$\text{grad } f(t\underline{x})\cdot t\underline{x} = \alpha f(t\underline{x})$$

da cui, ponendo $t=1$, si ricava la (1.3) .

Consideriamo ora un insieme di raggi uscenti dall'origine, <u>simmetrico</u> rispetto all'origine e supponiamo che la <u>(1.1)</u> sia verificata per ogni t <u>reale non nullo</u> . Diremo allora che f è <u>omogenea</u> di grado α in A .

Ad esempio, le funzioni considerate negli esempi 1 , 2 e 3 sono omogenee, rispettivamente di grado 2 , -2 , 0 nel loro insieme di definizione.

Le proprietà delle funzioni positivamente omogenee possono venire facilmente estese, con ovvie modifiche, alle funzioni omogenee.

Consideriamo ora in particolare un polinomio in $\underline{x}$ a coefficienti reali,omogeneo di grado k, cioè un polinomio $P(\underline{x})$ della forma

$$P(\underline{x}) = \sum_{\substack{i_1,i_2,\dots,i_n\geq 0\\ i_1+i_2+\dots+i_n=k}} a_{i_1,i_2,\dots,i_n}\, x_1^{i_1} x_2^{i_2}\cdots x_n^{i_n} \ .$$

E' facile verificare, tenendo conto dell'osservazione 2, che valgono le seguenti proprietà.

<u>Lemma</u> 1.1 - Se $P(\underline{x}^\circ)\neq 0$, allora $P(\underline{x})\neq 0$ in ogni punto della retta $(\underline{0},\underline{x}^\circ)$ diverso dall'origine e precisamente

a) se k è pari, $P(\underline{x})$ ha il segno di $P(\underline{x}^\circ)$ in ogni punto $\underline{x}\neq\underline{0}$ di tale retta;

b) se k è dispari, $P(\underline{x})$ ha il segno di $P(\underline{x}^\circ)$ nei punti del raggio $(\underline{0},\underline{x}^\circ)$ ed ha segno opposto nei punti del raggio $(\underline{0}, -\underline{x}^\circ)$.

<u>Lemma</u> 1.2 - Se $P(\underline{x})\neq 0$, allora al tendere di $\underline{x}$ a $\underline{0}$ sulla retta $(\underline{0},\underline{x}^\circ)$, $P(\underline{x})$ è infinitesimo di ordine k rispetto a $\|\underline{x}\|$.

<u>Lemma</u> 1.3 - Se P si annulla soltanto nell'origine, allora, per $\underline{x}\to\underline{0}$, $P(\underline{x})$ è infinitesimo di ordine k rispetto a $\|\underline{x}\|$.

Infatti, posto

$$m = \operatorname{Min}_{\|\underline{x}\|=1} |P(\underline{x})| \quad , \quad M = \operatorname{Max}_{\|\underline{x}\|=1} |P(\underline{x})|$$

risulta,

$$m\|\underline{x}\|^k \leq |P(\underline{x})| \leq M\|\underline{x}\|^k \quad , \forall \underline{x} \in R^n .$$

Poichè $0 < m < M < +\infty$, ne segue l'asserto.

2. <u>Estremi liberi delle funzioni di più variabili : condizioni necessarie</u>.

Consideriamo una funzione reale di n variabili definita in un aperto $A \subset R^n$ $(n \geq 1)$. Il problema della ricerca degli estremi (massimi e minimi) relativi di f in A viene abitualmente chiamato (per motivi che vedremo in seguito) problema della ricerca degli estremi <u>liberi</u> di f. Esso presenta nel caso $n > 1$ aspetti simili a quelli dell'analoga ricerca nel caso $n=1$. Ovviamente, se $n > 1$ gli sviluppi analitici risultano più complessi; tuttavia dal punto di vista concettuale i procedimenti seguiti sono gli stessi e i risultati che si ottengono generalizzano quelli analoghi del caso $n=1$.
Cominciamo con l'assegnare alcune condizioni necessarie per l'esistenza di un punto estremante per f in A.

<u>Teorema</u> 2.1 - Se f ha un estremo in un punto $\underline{x}^\circ \in A$ ed è differenziabile in tale punto, allora risulta:

$$(2.1) \qquad d\, f(\underline{x}^\circ;\Delta\underline{x}) = 0 \qquad \forall \Delta\underline{x} \in R^n .$$

Infatti, osserviamo preventivamente che

$$(2.2) \qquad f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ) = df(\underline{x}^\circ)+o(\|\Delta\underline{x}\|) .$$

Se non vale (2.1), per il Lemma 1.1 deve essere $df(\underline{x}^\circ;\Delta\underline{x})\neq 0$ in un insieme $I\neq\emptyset$ di rette uscenti dall'origine (ad eccezione del punto $\Delta\underline{x}=\underline{0}$) e, per il Lemma 1.2, al tendere di $\Delta\underline{x}$ a $\underline{0}$ in tale insieme, è infinitesimo del primo ordine. Ma allora da (2.2) segue che per $\Delta\underline{x}\to\underline{0}$ in I risulta

$$f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=df(\underline{x}^\circ;\Delta\underline{x})(1+o(1)) .$$

Ma per il lemma 1.1 (punto b)) $df(\underline{x}^\circ)$ non può avere segno costante al variare di $\Delta\underline{x}$ in alcun intorno dell'origine in I. Quindi $f(\underline{x})-f(\underline{x}^\circ)$ non ha segno costante in alcun intorno di $\underline{x}^\circ$, il che è assurdo.

<u>Corollario</u> 2.1 - Se f è differenziabile in A, i suoi eventuali punti estremanti soddisfano la condizione

$$(2.3) \qquad d\, f(\underline{x};\Delta\underline{x}) = 0 \qquad \forall \Delta\underline{x} \in R^n .$$

<u>Osservazioni</u> - 1) Se la (2.3) è soddisfatta per $\underline{x}=\underline{x}^\circ$, la superficie rappresentativa di f in $R^n \times R$ ha nel punto $\underline{x}^\circ$ piano tangente orizzontale. Ciò non implica necessariamente che $\underline{x}^\circ$ sia un punto di massimo o di minimo. I punti nei quali risulta soddisfatta la (2.3) senza che f abbia ivi un estremo vengono chiamati <u>punti di sella</u>, oppure <u>colli</u>. Tutti i punti nei quali vale (2.3) vengono chiamati <u>punti stazionari</u> di f.

2) La (2.3) equivale al sistema di n equazioni nelle n incognite $x_1,\dots,x_n$:

$$(2.4) \qquad \frac{\partial f}{\partial x_i}(x_1,\dots,x_n) = 0 \qquad i=1,\dots,n .$$

Se f è differenziabile in A, le soluzioni del sistema (2.4) permettono di individuare gli (eventuali) punti estremanti di f.

Esempi

1) Sia $f: R^3 \to R$ definita da $f(x_1,x_2,x_3)=x_1+e^{x_1}+x_2^2+x_3^2$. Il sistema (2.4) diviene

$$\begin{cases} 1+e^{x_1}=0 \\ 2x_2=0 \\ 3x_3=0 \end{cases}$$

e non ha alcuna soluzione. Quindi f non ha punti estremanti in R^3 .

2) Sia $f: R^3 \to R$ definita da $f(x_1,x_2,x_3)=x_1x_2+x_3^2$. Il sistema (2.4) diviene

$$\begin{cases} x_2=0 \\ x_1=0 \\ 2x_3=0 \end{cases}$$

ed ha solo la soluzione nulla. L'unico eventuale punto estremante per f è quindi l'origine. Si osservi però che f non ha segno costante in alcun intorno dell'origine: infatti p.e. se $x_3=0$ e x_1 e x_2 sono concordi di segno, f è positiva, mentre se $x_3=0$ e x_1 e x_2 sono discordi di segno, f è negativa. L'origine è quindi un punto di sella per f e non un punto estremante.

3) Sia $f: R^3 \to R$ definita da $f(x_1,x_2,x_3)=x_2^2(x_1-e^{x_1})^2+x_2^2+x_3^2$. Il sistema (2.4) diviene :

$$\begin{cases} 2x_2^2(x_1-e^{x_1})(1-e^{x_1})=0 \\ 2x_2\cdot((x_1-e^{x_1^2})+1)=0 \\ 2x_3=0 \end{cases}$$

e ha le soluzioni (c,o,o)(c reale qualsiasi). Si osservi ora che f in questi punti si annulla e non assume mai valori negativi. Ne segue che tutti i punti dell'asse x_1 sono per f punti di minimo (assoluto e debole).

Il teorema 2.1 rientra come caso particolare nel seguente

Teorema 2.2 - La funzione f abbia un estremo in un punto $\underline{x}^\circ \in A$ e sia k volte differenziabile in tale punto. Inoltre sia

$$(2.5) \qquad d\,f(\underline{x}^\circ) = d^2f(\underline{x}^\circ)=\ldots=d^{k-1}f(\underline{x}^\circ) = 0$$

mentre il differenziale k-mo non sia identicamente nullo. Allora k è pari e inoltre, per ogni $\Delta\underline{x} \in R^n$, risulta

$$(2.6) \qquad \begin{array}{ll} d^kf(\underline{x}^\circ;\Delta\underline{x})\le 0 & \text{se } \underline{x}^\circ \text{ è un punto di massimo} \\ d^kf(\underline{x}^\circ;\Delta\underline{x})\ge 0 & \text{se } \underline{x}^\circ \text{ è un punto di minimo.} \end{array}$$

Infatti, per il teorema 11.1 del cap.II, si ha, tenuto conto di (2.5):

$$f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)= \frac{1}{k!}\, d^kf(\underline{x}^\circ;\Delta\underline{x})+o(\|\Delta\underline{x}\|^k) \quad .$$

Se $d^kf(\underline{x}^\circ;\Delta\underline{x})$ non è identicamente nullo, esso, per il Lemma 1.1, è diverso da zero al variare di $\Delta\underline{x}$ in un insieme $I\neq\emptyset$ di rette uscenti dall'origine (private del punto $\Delta\underline{x}=\underline{0}$) e al tendere di $\Delta\underline{x}$ a $\underline{0}$ in I, per il Lemma 1.2, è infinitesimo di ordine k . Quindi, al tendere di $\Delta\underline{x}$ a $\underline{0}$ in I risulta

$$(2.7) \qquad f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=d^kf(\underline{x}^\circ;\Delta\underline{x})(1+o(1)) \quad .$$

Allora, per $\Delta\underline{x} \in I$ abbastanza piccolo, $d^kf(\underline{x}^\circ;\Delta x)$ ha il segno di $f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)$ e quindi risulta

$$(2.8) \qquad \begin{array}{ll} d^kf(\underline{x}^\circ) < 0 & \text{se } \underline{x}^\circ \text{ è di massimo} \\ d^kf(\underline{x}^\circ) > 0 & \text{se } \underline{x}^\circ \text{ è di minimo.} \end{array}$$

Di conseguenza, per il Lemma 1.1, k deve essere pari e le (2.8) devono essere soddisfatte per ogni $\Delta\underline{x} \in I$. Poichè inoltre $d^kf(\underline{x}^\circ;\Delta\underline{x})=0$ se $\underline{x}\notin I$, le (2.6) valgono per ogni $\Delta\underline{x}\in R^n$ e il teorema è dimostrato.

Osservazione - Le disuguaglianze (2.6) non valgono, in generale, in senso forte per ogni $\Delta\underline{x}\neq\underline{0}$, neanche se $\underline{x}^\circ$ è un punto di massimo forte o di minimo forte.

Si consideri infatti, come contro-esempio, la funzione $f: R^2 \to R$ definita da $f(x_1,x_2)=x_1^2+x_2^4$. Per tale funzione l'origine è un punto di minimo forte; $d\,f(\underline{0})=0$; $d^2f(\underline{0}) = 2\Delta x_1^2$ e quindi $d^2f(\underline{0})$ non soddisfa la condizione $d^2f(\underline{0})>0$ per $\Delta\underline{x}\neq\underline{0}$.

3. Estremi liberi delle funzioni di più variabili : condizioni sufficienti.

Anche in questo paragrafo consideriamo una funzione f reale definita in un aperto $A \subset R^n$ $(n \geq 1)$. Vale il

Teorema 3.1 - La funzione f sia k volte differenziabile in un punto $\underline{x}^\circ \in A$ e sia

(3.1) $$d\, f(\underline{x}^\circ)=d^2 f(\underline{x}^\circ)=\ldots=d^{(k-1)} f(\underline{x}^\circ)=0 \quad .$$

Se inoltre

(3.2) $$d^k f(\underline{x}^\circ;\Delta\underline{x}) < 0 \qquad \forall\, \Delta\underline{x} \neq \underline{0} ,$$

allora $\underline{x}^\circ$ è un punto di massimo forte; se

(3.3) $$d^k f(\underline{x}^\circ;\Delta\underline{x}) > 0 \qquad \forall\, \Delta\underline{x} \neq \underline{0} ,$$

$\underline{x}^\circ$ è un punto di minimo forte.

Infatti, se vale (3.1), per il Teorema 11.1 del capitolo II si ha

$$f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=\frac{1}{k!}\, d^k f(\underline{x}^\circ;\Delta\underline{x})+o(\|\Delta\underline{x}\|^k) \quad .$$

Inoltre, se vale (3.2) oppure (3.3), dal Lemma 1.3 si ricava per $\Delta\underline{x}\to\underline{0}$

$$f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=\frac{1}{k!}\, d^k f(\underline{x}^\circ;\Delta\underline{x})(1+o(1)) \quad .$$

Quindi, se $\|\Delta\underline{x}\|$ è abbastanza piccola, $f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)$ ha il segno di $d^k f(\underline{x}^\circ;\Delta\underline{x})$ e il teorema è dimostrato.

Osservazioni - 1) Affinchè valga (3.2) oppure (3.3), per il Lemma 1.1 è necessario che k sia pari.

2) Le condizioni sufficienti (3.2) e (3.3) sono più forti delle corrispondenti condizioni necessarie (2.6).

3) Le condizioni (2.6) non sono sufficienti per assicurare l'esistenza di un estremo nel punto $\underline{x}^\circ$.

Si consideri infatti, come contro-esempio, la funzione $f: R^2 \to R$ definita da $f(x_1,x_2)=x_2^2-x_1^4$, per la quale risulta $d\,f(\underline{0})=0$, $d^2f(\underline{0})=2\Delta x_2^2 \geq 0$ $\forall \Delta\underline{x}$. Eppure $f(x_1,x_2)=(x_2-x_1^2)(x_2+x_1^2)$ non ha un estremo nell'origine .

4) Se $n=1$, il teorema 3.1 coincide con una ben nota condizione sufficiente affinchè una funzione reale di una variabile reale abbia un estremo in x°.

Dai teoremi 2.2 e 3.1 si ricava come corollario il seguente

Teorema 3.2 - La funzione f sia k volte differenziabile in un punto $\underline{x}^\circ \in A$ e sia $df(\underline{x}^\circ)=d^2f(\underline{x}^\circ)=\ldots=d^{k-1}f(\underline{x}^\circ)=0$, mentre $d^k f(\underline{x}^\circ)$ non sia identicamente nullo. Allora se k è dispari, $\underline{x}^\circ$ è un colle. Se k è pari, allora

a) se $d^k f(\underline{x}^\circ;\Delta\underline{x})<0$ per ogni $\Delta\underline{x}\neq\underline{0}$, $\underline{x}^\circ$ è un punto di massimo forte;

b) se $d^k f(\underline{x}^\circ;\Delta\underline{x})>0$ per ogni $\Delta\underline{x}\neq\underline{0}$, $\underline{x}^\circ$ è un punto di minimo forte;

c) se $d^k f(\underline{x}^\circ;\Delta\underline{x})$ non ha segno costante, $\underline{x}^\circ$ è un colle.

Esempi

1) Consideriamo la funzione $f: R^3 \to R$ definita da $f(x,y,z)=x-e^x+y^2+z^2$. I suoi eventuali punti estremanti vanno ricercati fra le soluzioni del sistema (2.4)

$$\begin{cases} \dfrac{\partial f}{\partial x} = 1-e^x=0 \\ \dfrac{\partial f}{\partial y} = 2y = 0 \\ \dfrac{\partial f}{\partial z} = 2z = 0 \end{cases}$$

che ha soltanto la soluzione nulla. Poichè $\dfrac{\partial^2 f}{\partial x^2}=-e^x$, $\dfrac{\partial^2 f}{\partial y^2}=\dfrac{\partial^2 f}{\partial z^2}=2$ e le derivate seconde miste sono identicamente nulle , si ha

$$d^2f(\underline{0}) = -(\Delta x)^2+2(\Delta y)^2+2(\Delta z)^2.$$

Poichè $d^2f(\underline{0})$ non ha evidentemente segno costante, l'origine è un colle.
Allo stesso risultato si può anche giungere osservando che in un in

torno dell'origine si ha

$$f(x,y,z)-f(0,0,0) = x-\left(1+\frac{x}{1!}+\frac{x^2}{2!}+\dots\right)+y^2+z^2+1=$$

$$= -\frac{1}{2}x^2+y^2+z^2+o(x^2+y^2+z^2)$$

e che il polinomio $-\frac{1}{2}x^2+y^2+z^2$ non ha segno costante in alcun intorno dell'origine.

2) Consideriamo la funzione $f: R^3 \to R$ definita da

$$f(x,y,z)=(x^2+1+yz(e^z+1))^{-1/2}.$$

Osserviamo preventivamente che i suoi eventuali massimi e minimi sono, rispettivamente, i minimi e i massimi positivi della funzione

$$g(x,y,z) = x^2+1+yz(e^z+1) .$$

Per la funzione g il sistema (2.4) diviene

$$\begin{cases} \dfrac{\partial g}{\partial x} = 2x = 0 \\ \dfrac{\partial g}{\partial y} = z(e^z+1) = 0 \\ \dfrac{\partial g}{\partial z} = y(e^z+1+ze^z) = 0 \end{cases}$$

ed ha soltanto la soluzione nulla. Poichè $g(\underline{0})>0$, se g ha un estremo nell'origine anche f lo ha. Calcoliamo d^2g. Risulta

$$\frac{\partial^2 g}{\partial x^2}=2 , \quad \frac{\partial^2 g}{\partial y^2}=0 , \quad \frac{\partial^2 g}{\partial z^2}=y(2e^z+ze^z) , \quad \frac{\partial^2 g}{\partial x\partial y}=\frac{\partial^2 g}{\partial x\partial z}=0 \quad e$$

$$\frac{\partial^2 g}{\partial y\partial z}=e^z+1+z\,e^z \quad \text{e quindi} \quad d^2g(\underline{0}) = 2(\Delta x)^2+4\,\Delta y\,\Delta z .$$

Poichè $d^2g(\underline{0})$ non ha segno costante, g non ha alcun estremo e quindi nemmeno f ne ha.
Al risultato si poteva anche giungere osservando che in un intorno dell'origine risulta

$$g(x,y,z)-g(0,0,0) = x^2+2yz+o(x^2+y^2+z^2)$$

e che il polinomio x^2+2yz non ha segno costante in un intorno dell'origine.

3) Determiniamo gli eventuali estremi di $f(x,y,z)= x \operatorname{sen} x + y^2+z^2(4+y^2)$ nella sfera $x^2+y^2+z^2<1$.
Tali punti vanno ricercati fra le soluzioni del sistema

$$\begin{cases} \dfrac{\partial f}{\partial x} = \operatorname{sen} x + x\cos x = 0 \\ \dfrac{\partial f}{\partial y} = 2y(1+z^2) = 0 \\ \dfrac{\partial f}{\partial z} = 2z(4+y^2) = 0 \end{cases} .$$

Poichè l'equazione $\operatorname{sen} x + x\cos x = 0$ ha come unica soluzione in modulo minore di uno la soluzione nulla, il sistema ha, nella sfera unitaria, soltanto la soluzione nulla. Con facili calcoli si trova:

$$d^2f(\underline{0}) = 2(\Delta x)^2+2(\Delta y)^2+8(\Delta z)^2$$

e quindi l'origine è un punto di minimo.
Al risultato si può anche giungere osservando che per $(x,y,z)\to\underline{0}$ risulta

$$f(x,y,z)-f(0,0,0) = x^2+y^2+4z^2+o(x^2+y^2+z^2) .$$

Prendiamo ora in esame il caso (non considerato dal Teorema 3.2) in cui per $\underline{x}=\underline{x}^\circ$ sia $df=d^2f=\dots=d^{k-1}f=0$ e inoltre d^kf abbia segno costante ma si annulli per qualche $\Delta\underline{x}\neq\underline{0}$.
Questo caso viene chiamato "caso ambiguo" perché non si può giungere ad alcuna conclusione senza procedere ad ulteriori analisi.
Il procedimento più semplice è quello di studiare direttamente il segno di $f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)$, almeno per $\Delta\underline{x}$ abbastanza piccolo.
Se questo non è possibile e f è differenziabile in $\underline{x}^\circ$ un sufficiente numero di volte, si può procedere come segue. Sia I l'insieme di rette di R^n sulle quali $d^kf(\underline{x}^\circ)=0$. Se $d^{k+1}f(\underline{x}^\circ)\neq0$ in qualche punto di I, $\underline{x}^\circ$ non può essere un punto estremante(°). Se invece
→ questa osservazione non vale in generale per studiare il [illegible]

$$d^{k+1}f(\underline{x}^\circ)=0 \quad e \quad d^{k+2}f(\underline{x}^\circ)\neq0 \quad \forall\Delta x\in I, \Delta\underline{x}\neq0,$$

[illegible] f non diff

si dovrà esaminare il segno (in R^n!) di

$\Delta f = y^2 + x^2|x| - x^2y$

$$\Delta f=\frac{1}{k!}\,d^kf(\underline{x}^\circ)+\frac{1}{(k+1)!}\,d^{k+1}f(\underline{x}^\circ)+\frac{1}{(k+2)!}\,d^{k+2}f(\underline{x}^\circ) \quad (^*)$$

e così via.

(°) Infatti k+1 è dispari e $d^{k+1}f(\underline{x}^\circ)$ non può avere segno costante in I.
(*) Si noti che non basta esaminare il segno di $d^{k+2}f(\underline{x}^\circ)$ in I (si veda il successivo esempio 2).

Esempi

1) Consideriamo la funzione

$$f(x,y,z) = x^2y + xy^2 + \text{sen}^3 z \quad .$$

I suoi eventuali punti estremanti vanno ricercati fra le soluzioni del sistema

$$\begin{cases} \dfrac{\partial f}{\partial x} = 2xy + y^2 = 0 \\ \dfrac{\partial f}{\partial y} = x^2 + 2xy = 0 \\ \dfrac{\partial f}{\partial z} = 3\,\text{sen}^2 z \cos z = 0 \end{cases}$$

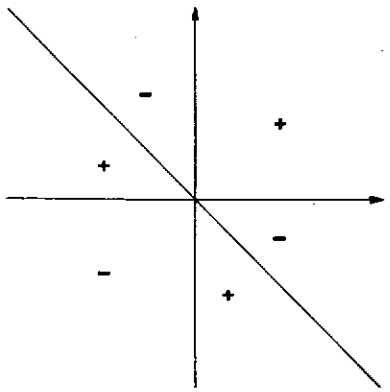
e quindi fra i punti $P_k=(0,0,k\pi)$ e $Q_k=(0,0,(\frac{1}{2}+k)\pi)$ (k intero relativo). Nei punti P_k tutte le derivate parziali seconde sono nulle, mentre p. e. $\dfrac{\partial^3 f}{\partial x^2 \partial y} = 2$.

Quindi si ha

$$d^2f(0,0,k\pi) \equiv 0 \quad , \quad d^3f(0,0,k\pi) \not\equiv 0$$

e di conseguenza i punti P_k non sono punti estremanti.
Nei punti Q_k le derivate parziali seconde sono nulle ad eccezione di $\dfrac{\partial^2 f}{\partial z^2}$ che assume il valore $-3\cdot(-1)^k=(-1)^{k-1}\cdot 3$. Allora

$$d^2f(0,0,(\tfrac{1}{2}+k)\pi) = (-1)^{k-1}\cdot 3\,(\Delta z)^2 \quad .$$



Quindi d^2f ha segno costante, ma si annulla se $\Delta z=0$ qualunque siano Δx e Δy. Si ha perciò il caso ambiguo: per decidere conviene esaminare direttamente il comportamento della funzione in un intorno dei punti Q_k. Osserviamo allora che

$$f(x,y,(\tfrac{1}{2}+k)\pi)-f(0,0,(\tfrac{1}{2}+k)\pi) = xy\,(x+y)$$

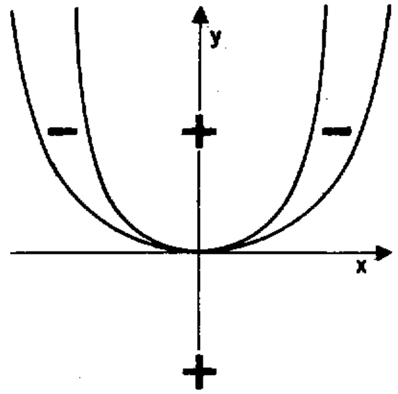
non ha evidentemente segno costante in un intorno di $x=y=0$ (vedi figura). Quindi Δf assume, in qualsiasi intorno del punto $(0,0,(\frac{1}{2}+k)\pi)$ valori di segno opposto. Neanche i punti Q_k sono perciò punti estremanti.

2) Consideriamo la funzione

$$f(x,y) = y^2 - 3x^2y + 2x^4 + 1 \quad .$$

Il sistema

$$\begin{cases} \dfrac{\partial f}{\partial x} = -6xy + 8x^3 = 0 \\ \dfrac{\partial f}{\partial y} = 2y - 3x^2 = 0 \end{cases}$$



ha l'unica soluzione $x=y=0$ e in tale punto $d^2f=2\Delta y^2$ ha segno costante, ma si annulla se $\Delta y=0$. Si ha quindi il caso ambiguo. Osserviamo allora che è

$$\Delta f=f(x,y)-f(0,0)=y^2-3x^2y+2x^4= (y-x^2)(y-2x^2) \quad .$$

Poichè (vedi figura) in qualunque intorno dell'origine Δf assume valori di segno contrario, l'origine non è un estremo per f, pur essendo un punto di minimo per qualsiasi sezione di $z=f(x,y)$ con un piano verticale contenente l'asse z .

Si noti che nel punto (0,0) risulta $d^2f=d^3f=0$ se $\Delta y=0$; $d^4f\geq 0$ per ogni $(\Delta x,\Delta y)\neq(0,0)$ e, in particolare, $d^4f>0$ per ogni $\Delta x\neq 0$; eppure l'origine non è un punto estremante.

Nel caso particolare k=2 il teorema 3.2 diviene

<u>Teorema</u> 3.3 - La funzione f sia due volte differenziabile in un punto $\underline{x}°\in A$ con $df(\underline{x}°)=0$ e $d^2f(\underline{x}°)$ non identicamente nullo. Allora, se la forma quadratica

$$Q(\underline{x})=\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}°)\,x_i x_j$$

è definita negativa, $\underline{x}_0$ è un punto di massimo forte; se Q è definita positiva, $\underline{x}°$ è un punto di minimo forte; se Q è indefinita, $\underline{x}°$ non è punto estremante; se infine Q è semidefinita si ha il caso ambiguo.

Infatti, basta osservare che $Q(\Delta\underline{x})=d^2f(\underline{x}°;\Delta\underline{x})$.

Più in particolare ancora, nel caso $n=2$, $k=2$ si ottiene il

<u>Teorema</u> 3.4 - Sia $n=2$; la funzione f sia due volte differenziabile in un punto $\underline{x}°\in A$ e sia $df(\underline{x}°)=0$. Allora, se il determinante hessiano di f nel punto $\underline{x}°$

$$H = \det\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}°)\right]_{i,j=1,\dots,n}$$

è positivo, $\underline{x}^\circ$ è un punto estremante e precisamente un massimo forte se $\frac{\partial^2 f}{\partial x_1^2}(\underline{x}^\circ) < 0$, un minimo forte se $\frac{\partial^2 f}{\partial x_1^2}(\underline{x}^\circ) > 0$; se $H<0$, $\underline{x}^\circ$ non è un punto estremante; se infine $H=0$ si ha il caso ambiguo.

Infatti, in questo caso la forma quadratica $Q(\underline{x})$ associata al punto $\underline{x}^\circ$ è data da

$$Q(\underline{x})= \frac{\partial^2 f}{\partial x_1^2}(\underline{x}^\circ)x_1^2+2\frac{\partial^2 f}{\partial x_1 \partial x_2}(\underline{x}^\circ)x_1x_2+\frac{\partial^2 f}{\partial x_2^2}(\underline{x}^\circ)x_2^2$$

e il suo discriminante è $-4H$. Dalle proprietà elementari dei polinomi di secondo grado e dal teorema 3.3 segue facilmente l'asserto.

E s e m p i

1) Consideriamo la funzione $f(x,y) = xy\, e^{-(x+y)}$. I suoi eventuali punti estremanti sono le soluzioni del sistema

$$\begin{cases} \frac{\partial f}{\partial x} = y(1-x)e^{-(x+y)} = 0 \\ \frac{\partial f}{\partial y} = x(1-y)e^{-(x+y)} = 0 \end{cases}$$

e cioè i punti $(0,0)$ e $(1,1)$. Poichè f non ha segno costante in un intorno dell'origine , tale punto non può essere un punto estremante e basta esaminare il punto $(1,1)$. Poichè in tale punto si ha

$$H = \begin{vmatrix} -e^{-2} & 0 \\ 0 & -e^{-2} \end{vmatrix} > 0 \quad e \quad \frac{\partial^2 f}{\partial x^2} = -e^{-2} < 0 ,$$ il punto $(1,1)$ è un punto di massimo(forte).

2) Consideriamo la funzione $f(x,y) = x^2 y\, e^{-(x^2+y^2)}$. I suoi punti estremanti vanno ricercati fra le soluzioni del sistema

$$\begin{cases} \frac{\partial f}{\partial x} = 2xy(1-x^2)e^{-(x^2+y^2)} = 0 \\ \frac{\partial f}{\partial y} = x^2(1-2y^2)e^{-(x^2+y^2)} = 0 \end{cases}$$

cioè nei punti $(0,k)$ (k reale qualunque), $(1, \frac{1}{\sqrt2})$, $(1,-\frac{1}{\sqrt2})$ $(-1, \frac{1}{\sqrt2})$ $(-1, -\frac{1}{\sqrt2})$. Un esame del segno di f mostra immediatamente che i punti $(0,k)$ sono minimi deboli se $k>0$, massimi deboli se $k<0$ e che il punto $(0,0)$ non è punto estremante. Nel punto $(1, \frac{1}{\sqrt2})$ risulta $H=8e^{-3}>0$ $\frac{\partial^2 f}{\partial x^2} = -2\sqrt2\, e^{-3/2} < 0$ e quindi tale punto è un massimo forte. Per motivi di simmetria allora anche il punto $(-1, \frac{1}{\sqrt2})$ è di massimo forte e i punti $(\pm1, -\frac{1}{\sqrt2})$ sono punti di minimo forte.

4. Funzioni implicite.

Consideriamo, a titolo di esempio, una funzione reale delle due variabili reali x,y che goda della seguente proprietà: fissato comunque x in R , l'equazione

$$(4.1) \qquad f(x,y) = 0$$

abbia una e una sola soluzione $y\in R$. Al variare del punto x fissato, la corrispondente soluzione y di (4.1) varia definendo così una funzione $y:R\to R$, $y=y(x)$, caratterizzata dal fatto che

$$f(x,y(x)) = 0 \qquad \forall x \in R .$$

E' evidente che tale funzione y risulta assegnata in via indiretta tramite l'equazione $f(x,y)=0$. Questa considerazione ci conduce alla segunte definizione di "funzione definita implicitamente dall'equazione $f(x,y)=0$ " (o anche "funzione implicita nell'equazione $f(x,y)=0$").
Siano A e B due insiemi di punti, rispettivamente,di R^n e R $(n\geq 1)$; sia f una funzione reale definita in $A\times B$ e si consideri l'equazione $f(\underline{x},y)=0$. Se esiste una funzione $y:R^n\to R$, $y=y(\underline{x})$, definita in un sottoinsieme $A^*\subset A$ a valori in B e tale che risulti

$$f(\underline{x}, y(\underline{x})) = 0 \qquad \forall \underline{x} \in A^* ,$$

diremo che y è <u>definita implicitamente dall'equazione</u> $f(\underline{x},y)=0$.

E s e m p i

1) L'equazione $(x^2+1)y+x^2-1=0$ $(x,y\in R)$ definisce implicitamente una e una sola funzione reale della variabile reale x, cioè la funzione

$$y(x) = \frac{1-x^2}{1+x^2} \qquad \forall x\in R .$$

2) L'equazione $xy-1=0$ $(x,y\in R)$ definisce implicitamente una e una sola

funzione reale di x in $R-\{0\}$, cioè la funzione

$$y(x) = \frac{1}{x} .$$

3) L'equazione $e^y - x = 0$ $(x,y \in R)$ definisce implicitamente una e una sola funzione reale di x in $(0,+\infty)$ e cioè la funzione

$$y(x) = \log x .$$

4) L'equazione $z\, e^{xy} - 1 = 0$ $(x,y,z \in R)$ definisce implicitamente una e una sola funzione reale delle due variabili reali x e y, cioè la funzione

$$z(x,y) = e^{-xy} \qquad \forall (x,y) \in R^2$$

ed anche una e una sola funzione y reale di (x,z) in $(R-\{0\}) \times R^+$, cioè la funzione

$$y(x,z) = -\frac{1}{x} \log z .$$

La definizione di funzione implicita può venire estesa in modo ovvio anche al caso di funzioni vettoriali di un vettore. Si ottiene così la seguente definizione.

Siano A e B due insiemi di punti, rispettivamente, di R^n e R^m $(n,m \geq 1)$; sia $\underline{f}$ una funzione definita in $A \times B$ a valori in R^m e si consideri l'equazione

$$\underline{f}(\underline{x},\underline{y}) = \underline{0} . \tag{4.2}$$

Se esiste una funzione $\underline{y} = \underline{y}(\underline{x})$ definita in un sottoinsieme $A^* \subset A$, a valori in B e tale che

$$\underline{f}(\underline{x},\underline{y}(\underline{x})) = \underline{0} \qquad \forall\, \underline{x} \in A^*,$$

diremo che $\underline{y}$ è definita implicitamente dall'equazione $\underline{f}(\underline{x},\underline{y}) = \underline{0}$.

Si noti che se $m > 1$ l'equazione (vettoriale!) (4.2) equivale al sistema di m equazioni scalari nelle n+m componenti dei vettori $\underline{x}$ e $\underline{y}$

$$\begin{cases} f_1(x_1,\dots,x_n\,;\, y_1,\dots,y_m) = 0 \\ f_2(x_1,\dots,x_n\,;\, y_1,\dots,y_m) = 0 \\ \dots\dots\dots \\ f_m(x_1,\dots,x_n\,;\, y_1,\dots,y_m) = 0 \end{cases} .$$

Terminiamo con le seguenti osservazioni.

1) L'equazione (4.2) non definisce sempre necessariamente una funzione implicita.

Basta considerare per esempio l'equazione $x^2 + y^2 + 1 = 0$ $(x,y \in R)$.

2) Una funzione implicita nell'equazione (4.2) non è sempre necessariamente definita per ogni $\underline{x} \in A$.

Basta considerare gli esempi 2 e 3 di questo paragrafo.

3) L'esistenza o meno di una funzione definita implicitamente da (4.2) dipende in generale, da A e B.

Infatti, si consideri l'equazione $x + y^2 = 0$. Se $A = R^+$ e $B = R^+$ tale equazione non definisce implicitamente alcuna funzione; se $A = R^-$ e $B = R^+$ essa definisce implicitamente la funzione $y(x) = \sqrt{-x}$ $\forall x \in R^-$.

4) Anche l'unicità o meno della funzione implicita dipende, in generale, da A e B.

Infatti, si consideri l'equazione

$$x^2 + (y^2+y)x + y^3 = (x+y)(x+y^2) = 0 . \tag{4.3}$$

Se $A = R^+$ e $B = R$ essa definisce implicitamente una e una sola funzione di x, cioè la funzione $y(x) = -x$. Se $A = R^-$ e $B = R^-$ essa definisce ancora una e una sola funzione implicita: la funzione $y(x) = -\sqrt{-x}$. Se invece $A = R^-$ e $B = R^+$, essa definisce implicitamente almeno due funzioni, e cioè quelle assegnate rispettivamente da

$$y_1(x) = -x \;,\quad y_2(x) = \sqrt{-x} \qquad \forall x \in R^- ;$$

anzi, ne definisce implicitamente infinite. Infatti, sia I un qualsiasi sottoinsieme di R^-; la funzione y_I definita da

$$y_I(x) = \begin{cases} -x & \forall x \in I \\ \sqrt{-x} & \forall x \in R^- - I \end{cases}$$

soddisfa evidentemente la (4.3) per ogni x di R^- e pertanto è definita implicitamente da tale equazione.

Una volta fissati A e B sorge allora il problema di assegnare condizioni sufficienti affinchè l'equazione (4.2) definisca una e una sola funzione implicita. I teoremi di questo tipo si dividono in due categorie: quelli che assicurano l'esistenza e l'unicità della funzione implicita a valori in B <u>per ogni $\underline{x} \in A$</u> (teoremi di esistenza e unicità "in grande") e quelli che assicurano l'esistenza e l'unicità <u>in</u>

<u>un opportuno intorno di un punto</u> $\underline{x}^\circ \in A$ della funzione implicita a valori in un opportuno intorno di un punto $\underline{y}^\circ \in B$ (teoremi di esistenza e unicità "in piccolo").

5. <u>Teorema di esistenza e unicità "in grande"</u>.

Prendiamo in esame soltanto il caso $m=1$. Vale il

Vale anche se $b_1 = b_1(x)$ e $b_2 = b_2(x)$

<u>Teorema</u> 5.1 - Siano: A un sottoinsieme di R^n $(n \geq 1)$; B un intervallo di R , $B=(b_1, b_2)$, $-\infty \leq b_1 < b_2 \leq +\infty$. La funzione f sia definita in $A \times B$ a valori in R e soddisfi le condizioni seguenti:

1) per ogni $\underline{x} \in A$ la funzione $f(\underline{x}, \cdot)$ sia continua e strettamente crescente in B ;

2) per ogni $\underline{x} \in A$ sia

$\lim_{y \to b_1^+} f(\underline{x}, y) < 0$, $\lim_{y \to b_2^-} f(\underline{x}, y) > 0$.

Allora esiste una e una sola funzione $y=y(\underline{x})$ definita in A a valori in B e soddisfacente l'equazione

$$f(\underline{x}, y(\underline{x})) = 0 \qquad \forall\, \underline{x} \in A . \tag{5.1}$$

Infatti, consideriamo per ogni $\underline{x} \in A$ la funzione ausiliaria $\varphi_{\underline{x}} = f(\underline{x}, \cdot)$. Per la condizione 2) essa ha in B estremo inferiore negativo e estremo superiore positivo e quindi, per il teorema di Weierstrass, assume in almeno un punto $y \in B$ il valore zero. Tale punto y è unico, perchè $\varphi_{\underline{x}}$ è strettamente crescente in B ; al variare di $\underline{x}$ in A esso varia con $\underline{x}$ definendo così l'unica funzione $y=y(\underline{x})$ a valori in B che soddisfa la (5.1).

<u>Osservazioni</u> - 1) Si può ovviamente supporre $f(\underline{x}, \cdot)$ anzichè strettamente crescente, strettamente decrescente in B per ogni $\underline{x} \in A$; in tal caso le 2) si modificano in modo ovvio.

2) Se $b_1 > -\infty$ e la condizione 1) è soddisfatta in $[b_1, b_2)$, la prima delle condizioni 2) può essere sostituita da $f(\underline{x}, b_1) \leq 0$ $\forall\, \underline{x} \in A$. Analogamente dicasi per b_2 .

3) La condizione 1) è certamente soddisfatta se f è continua in $A \times B$ ed esiste in ogni punto di $A \times B$, $\frac{\partial f}{\partial y}$ e tale derivata è sempre positiva.

<u>Esempi</u>

1) Consideriamo l'equazione

$$f(x,y) = e^{x^2 y} + y - x = 0 \qquad x , y \in R .$$

La funzione f è continua e $\frac{\partial f}{\partial y} = x^2 e^{x^2 y} + 1$ è positiva in $R \times R$. Inoltre $\lim_{y \to -\infty} f(x,y) = -\infty$ e $\lim_{y \to +\infty} f(x,y) = +\infty$ $\forall x \in R$. Per il teorema 5.1 l'equazione data definisce quindi nell'intervallo $(-\infty, +\infty)$ una e una sola funzione reale $y = y(x)$.

2) Consideriamo l'equazione

$$f(x,y) = e^y + e^{(x^2+1)y^3} + x^2 - 1 = 0 \qquad x, y \in R .$$

f è continua e $\frac{\partial f}{\partial y} = e^y + 3(x^2+1)y^2 e^{(x^2+1)y^3}$ è positiva in $R \times R$. Inoltre si ha

$\lim_{y \to -\infty} f(x,y) = x^2 - 1$, $\lim_{y \to +\infty} f(x,y) = +\infty$ $\forall x \in R$.

Poichè $x^2 - 1 \geq 0$ per $|x| \geq 1$ e $f(x, \cdot)$ è crescente, l'equazione data non può essere soddisfatta in alcun punto di $(-\infty, -1] \times R$ o di $[1, +\infty) \times R$ e quindi non può definire alcuna funzione implicita per $|x| \geq 1$. Invece in $(-1,1) \times R$ le condizioni del teorema 5.1 sono evidentemente soddisfatte e quindi l'equazione data definisce implicitamente nell'intervallo $(-1,1)$ (e solo in esso!) una funzione $y=y(x)$ e una solamente.

3) Consideriamo l'equazione

$$f(x,y) = x^4 + y^3 + x^2 + y^2 + 1 = 0 \qquad x, y \in R .$$

f è continua in $R \times R$; $\frac{\partial f}{\partial y} = 3y^2 + 2y = y(3y+2)$ si annulla sulle rette $y=0$ e $y = -\frac{2}{3}$. Conviene quindi considerare separatamente i tre insiemi di $R \times R$: $R \times R^+$, $R \times (-\frac{2}{3}, 0)$, $R \times (-\infty, -\frac{2}{3})$.

In $R \times R^+$ $\frac{\partial f}{\partial y}$ è positiva; $\lim_{y \to 0^+} f(x,y) = x^4 + x^2 + 1 > 0$ $\forall x \in R$ e quindi l'equazione data non può essere soddisfatta in tale insieme.

In $R \times (-\frac{2}{3}, 0)$ $\frac{\partial f}{\partial y}$ è negativa; $\lim_{y \to 0^-} f(x,y) > 0$ $\forall x \in R$ e quindi anche in questo insieme l'equazione data non ha alcuna soluzione.

Consideriamo infine l'insieme $R \times (-\infty, -\frac{2}{3})$. In esso $\frac{\partial f}{\partial y}$ è positiva;

$\lim\limits_{y\to-\infty} f(x,y)=-\infty$ e $\lim\limits_{y\to -\frac{2}{3}^-} f(x,y)=x^4+x^2+\frac{31}{27}>0 \quad \forall x\in R$ e quindi in $R\times(-\infty,-\frac{2}{3})$ vale il teorema 5.1.

Riassumendo i risultati ottenuti, possiamo concludere che l'equazione data definisce nell'intervallo $(-\infty,+\infty)$ una e una sola funzione implicita $y(x)$ e che tale funzione assume sempre valori minori di $-\frac{2}{3}$.

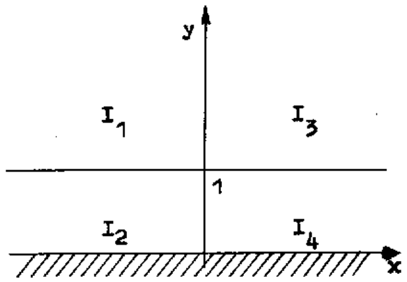
4) Consideriamo l'equazione

$$f(x,y) = y^2 - 2\log y + x^3 + x - 1 = 0 \qquad x\in R,\ y\in R^+.$$

f è continua in $R\times R^+$; $\frac{\partial f}{\partial y} = 2y - \frac{2}{y}$ si annulla per $y=1$.
Inoltre si ha

$\lim\limits_{y\to 0+} f(x,y)=+\infty$, $\lim\limits_{y\to 1\mp} f(x,y)=x(x^2+1)$,

$\lim\limits_{y\to+\infty} f(x,y)=+\infty$.



Conviene quindi considerare separatamente i quattro sottoinsiemi di $R\times R^+$ indicati in figura.
In I_3 e I_4 f non può avere zeri; quindi l'equazione data non definisce alcuna funzione implicita reale $y(x)$ per $x>0$.
In I_1 e in I_2 vale il teorema 5.1; quindi l'equazione data definisce per $x\leq 0$ una e una sola funzione implicita $y(x)$ a valori compresi fra 0 e 1 e una e una sola funzione implicita a valori maggiori di uno. Entrambe queste funzioni per $x=0$ assumono il valore uno.

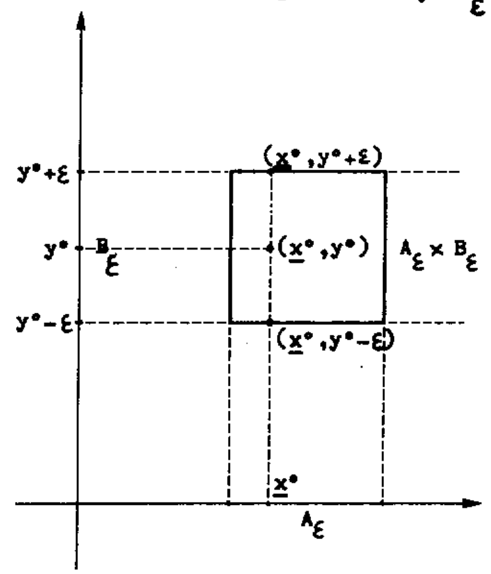
6. Continuità della funzione implicita.

Teorema 6.1 - Siano soddisfatte le ipotesi del teorema 5.1 e inoltre f sia continua in $A\times B$. Allora la funzione $y(\underline{x})$ definita implicitamente in A dall'equazione $f(\underline{x},y)=0$ è continua in A.

Infatti, sia $\underline{x}^\circ$ un qualsiasi punto di A e supponiamo che il punto $y^\circ=y(\underline{x}^\circ)$ sia interno a B (°). Sia $\varepsilon>0$ arbitrario e sia B_ε l'intervallo $[y^\circ-\varepsilon, y^\circ+\varepsilon]$.

(°) Se y° è uno degli estremi di B la dimostrazione si modifica in modo ovvio.

Se ε è abbastanza piccolo, $B_\varepsilon\subset B$. Poichè $f(\underline{x}^\circ,y^\circ)=0$, dalle ipotesi del teorema 5.1 segue $f(\underline{x}^\circ,y^\circ-\varepsilon)<0$, $f(\underline{x}^\circ,y^\circ+\varepsilon)>0$. Per la continuità di f in $A\times B$ deve allora esistere un intorno A_ε di $\underline{x}^\circ$ in A tale che



$f(\underline{x},y^\circ-\varepsilon)<0$,

$f(\underline{x},y^\circ+\varepsilon)>0$

$\forall \underline{x}\in A_\varepsilon$.

Per il teorema 5.1 applicato ad $A_\varepsilon\times B_\varepsilon$ la funzione implicita $y(\underline{x})$ a valori in B, definita in un sol modo in A dall'equazione data, risulta, per $\underline{x}\in A_\varepsilon$, a valori in B_ε e cioè

$$|y(\underline{x}) - y(\underline{x}^\circ)| \leq \varepsilon \qquad \forall \underline{x}\in A_\varepsilon \quad , \text{ c.d.d.}$$

E' facile verificare che tutte le funzioni implicite considerate negli esempi del paragrafo 5 sono continue. In particolare si può affermare che l'equazione considerata nell'esempio 5 definisce implicitamente per $x\leq 0$ due e due sole funzioni reali $y(x)$ continue. Non si può invece affermare che essa definisce implicitamente due e due sole funzioni reali $y(x)$. (Si ricordi l'esempio 4 del paragrafo 4).

7. Differenziabilità della funzione implicita.

Teorema 7.1 - Siano soddisfatte le ipotesi del teorema 5.1 e inoltre f sia differenziabile in ogni punto interno ad $A\times B$ con $\frac{\partial f}{\partial y}\neq 0$. Allora la funzione implicita $y=y(\underline{x})$ è differenziabile in ogni punto interno ad A e

$$(7.1)\qquad \operatorname{grad} y(\underline{x}) = \left(- \frac{\frac{\partial f}{\partial x_1}(\underline{x},y(\underline{x}))}{\frac{\partial f}{\partial y}(\underline{x},y(\underline{x}))}, \dots, - \frac{\frac{\partial f}{\partial x_n}(\underline{x},y(\underline{x}))}{\frac{\partial f}{\partial y}(\underline{x},y(\underline{x}))} \right) .$$

Infatti, denotiamo (per semplicità di scrittura) con $\underline{z}$ il punto generico di $R^n \times R$, $\underline{z}=(\underline{x},y)$ e con

$$P(\underline{z}) = \left[\frac{\partial f}{\partial x_i}(\underline{z}) \right]_{i=1,\dots,n} \qquad p(\underline{z}) = \frac{\partial f}{\partial y}(\underline{z}) .$$

Se $\underline{z}=(\underline{x},y)$ è un qualsiasi punto interno ad $A \times B$, si ha, per ogni $\Delta\underline{z}=(\Delta\underline{x},\Delta y)$ abbastanza piccolo:

$$(7.2)\qquad \begin{aligned} f(\underline{z}+\Delta\underline{z})-f(\underline{z}) &= \operatorname{grad} f(\underline{z})\Delta\underline{z}+o(\|\Delta\underline{z}\|) = \\ &= P(\underline{z})\Delta\underline{x}+p(\underline{z})\Delta y+o(\|\Delta\underline{z}\|) \end{aligned} .$$

Consideriamo ora, in particolare, un punto $\underline{z}^\circ$ interno ad $A \times B$ della forma $(\underline{x}^\circ, y(\underline{x}^\circ))$ e conveniamo di scegliere, per ogni $\Delta\underline{x}$ abbastanza piccolo, il valore di Δy, anzichè arbitrariamente, nel modo seguente:

$$(7.3)\qquad \Delta y = y(\underline{x}^\circ+\Delta\underline{x}) - y(\underline{x}^\circ) .$$

In tal caso il primo membro di (7.2) è nullo e quindi da (7.2) si ricava

$$(7.4)\qquad \underline{P}(\underline{z}^\circ)\Delta\underline{x} + p(\underline{z}^\circ)\Delta y = o(\|\Delta\underline{z}\|) .$$

Osserviamo ora che risulta necessariamente

$$(7.5)\qquad \Delta y = O(\|\Delta\underline{x}\|) .$$

? Infatti, se così non fosse, esisterebbe una successione $\{\Delta_n\underline{x}\} \to \underline{0}$ tale che i corrispondenti valori $\Delta_n y$ assegnati da (7.3) soddisferebbero la condizione $\frac{\Delta_n y}{\|\Delta_n\underline{x}\|} \longrightarrow \infty$. Ma allora da (7.4) (riferita a $\Delta_n\underline{x}$), dividendo per $\Delta_n y$ e ricordando il lemma 3.1 del cap.III, si otterrebbe:

$$p(\underline{z}^\circ) = o(1)$$

assurdo, perchè $p(\underline{z}^\circ) \neq 0$.
Di conseguenza, deve valere (7.5) e allora

$$\|\Delta\underline{z}\| = \left\{ \|\Delta\underline{x}\|^2 + (\Delta y)^2 \right\}^{\frac{1}{2}} = O(\|\Delta\underline{x}\|)$$

e da (7.4) si ottiene

$$\Delta y = y(\underline{x}^\circ+\Delta\underline{x})-y(\underline{x}^\circ) = - \frac{P(\underline{z}^\circ)}{p(\underline{z}^\circ)}\Delta\underline{x} + o(\|\Delta\underline{x}\|)$$

e il teorema è dimostrato.

<u>Osservazioni</u> - 1) Da (7.1) si ottiene, per ogni $\underline{x}$ interno ad A, nel caso $n>1$:

$$(7.6)\qquad \frac{\partial y}{\partial x_i}(\underline{x}) = - \frac{\frac{\partial f}{\partial x_i}(\underline{x},y(\underline{x}))}{\frac{\partial f}{\partial y}(\underline{x},y(\underline{x}))} \qquad (i=1,\dots,n)$$

e nel caso $n=1$:

$$(7.7)\qquad y'(x) = - \frac{\frac{\partial f}{\partial x}(x,y(x))}{\frac{\partial f}{\partial y}(x,y(x))} .$$

Osserviamo che tali formule si ricordano facilmente. Basta infatti differenziare la relazione $f(\underline{x},y)=0$. Si ricava così

$$\frac{\partial f}{\partial x_1}dx_1 + \dots + \frac{\partial f}{\partial x_n}dx_n + \frac{\partial f}{\partial y}dy = 0 .$$

Di qui si ottiene, se $n>1$

$$dy = \sum_{i=1}^{n} \frac{\partial y}{\partial x_i}dx_i = \sum_{i=1}^{n} - \frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial y}}dx_i$$

da cui le (7.6).
Nel caso particolare $n=1$ si ottiene

$$dy = y'(x)dx = - \frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}dx$$

e quindi la (7.7).

2) Se vale il teorema 7.1 e f è differenziabile con continuità internamente ad $A \times B$, la funzione

implicita $y(\underline{x})$ è differenziabile con continuità internamente ad A .

3) Per dimostrare che la funzione implicita $y(\underline{x})$ è differenziabile in un punto $\underline{x}°$ interno ad A abbiamo utilizzato esclusivamente l'ipotesi che f sia differenziabile con $\frac{\partial f}{\partial y} \neq 0$ nel punto $(\underline{x}°, y(\underline{x}°))$. Di conseguenza, vale anche il teorema seguente.

Teorema 7.2 - Sia $y=y(\underline{x})$ una funzione reale definita implicitamente in un insieme $A \subset R^n$ dall'equazione $f(\underline{x},y)=0$ e sia $\underline{x}°$ un punto interno ad A. Allora, se f è differenziabile nel punto $(\underline{x}°, y(\underline{x}°))$ con $\frac{\partial f}{\partial y}(\underline{x}°, y(\underline{x}°)) \neq 0$, la funzione implicita è differenziabile in $\underline{x}°$ e vale (7.1).

Esempi

1) La funzione $y(x)$ definita implicitamente (vedasi par.5, esempio 1) dall'equazione

$$e^{x^2 y} + y - x = 0 \qquad x,y \in R$$

è differenziabile con continuità su R ed è, per ogni $x \in R$

$$y'(x) = -\frac{2xy(x)e^{x^2 y(x)} - 1}{x^2 e^{x^2 y(x)} + 1} .$$

In particolare per $x=0$, essendo evidentemente $y(0)=-1$ si ha $y'(0)=1$.

2) L'equazione

$$f(x_1,x_2,y)=(x_1^2+x_2^2)e^{x_1^2 y}+y = 0 \qquad x_1,x_2,y \in R$$

definisce implicitamente in R^2 una e una sola funzione reale $y=y(x_1,x_2)$ (infatti è facile verificare che tutte le condizioni del teorema 5.1 sono soddisfatte in $R^2 \times R$). Tale funzione è inoltre di classe $\mathscr{C}^1(R^2)$ e si ha

$$\frac{\partial y}{\partial x_1}(x_1,x_2) = -\frac{2x_1(1+(x_1^2+x_2^2)y(x_1,x_2))e^{x_1^2 y(x_1,x_2)}}{x_1^2(x_1^2+x_2^2)e^{x_1^2 y(x_1,x_2)}+1} =$$

$$= \frac{2x_1 y(x_1,x_2)(1+(x_1^2+x_2^2)y(x_1,x_2))}{(x_1^2+x_2^2)(1-x_1^2 y(x_1,x_2))}$$

$$\frac{\partial y}{\partial x_2}(x_1,x_2) = -\frac{2x_2 e^{x_1^2 y(x_1,x_2)}}{x_1^2(x_1^2+x_2^2)e^{x_1^2 y(x_1,x_2)}+1} = \frac{2x_2 y(x_1,x_2)}{(x_1^2+x_2^2)(1-x_1^2 y(x_1,x_2))} .$$

In particolare, nel punto (0,1), essendo $y(0,1)=-1$ si ha $\frac{\partial y}{\partial x_1} = 0$, $\frac{\partial y}{\partial x_2} = -2$.

8. Differenziali di ordine superiore al primo delle funzioni implicite.

La funzione f soddisfi in A×B le ipotesi del teorema 5.1 e sia k volte differenziabile internamente ad A×B . Allora da (7.1) si ricava immediatamente che la funzione implicita $y=y(\underline{x})$ è k volte differenziabile internamente ad A . Inoltre, se f è di classe $\mathscr{C}^{(k)}$ internamente ad A×B , la funzione implicita è di classe $\mathscr{C}^{(k)}$ internamente ad A.
Per ottenere nel modo più rapido i differenziali (e le derivate parziali) di ordine superiore al primo della funzione implicita $y(\underline{x})$, conviene differenziare k volte la funzione composta $f(\underline{x},y(\underline{x}))$, osservare che tutti i differenziali ottenuti sono nulli e ricavare dalle k equazioni così ottenute rispettivamente dy , $d^2y, \ldots, d^k y$.
A titolo di esempio, calcoliamo $y''(x)$ nel caso m=n=1.Con ovvio significato dei simboli si ha

$$\begin{cases} f'_x + f'_y y' = 0 \\ f''_{xx} + 2f''_{xy} y' + f''_{yy} y'^2 + f'_y y'' = 0 \end{cases}$$

da cui

$$f''_{xx} f'^2_y - 2f''_{xy} f'_x f'_y + f''_{yy} f'^2_x + f'^3_y y'' = 0$$

e quindi

$$y'' = -\frac{f''_{xx} f'^2_y - 2f''_{xy} f'_x f'_y + f''_{yy} f'^2_x}{f'^3_y} .$$

Ad esempio, la funzione $y: R \to R$ definita implicitamente dall'equazione

$$e^{x^2 y} + y - x = 0$$

(vedasi par.5, es.1) è di classe $\mathcal{C}^\infty(R)$. Le sue derivate prime e seconde soddisfano il sistema

$$\begin{cases} (2xy + x^2 y')e^{x^2 y} + y' - 1 = 0 \\ (2y + 4xy' + x^2 y'' + (2xy + x^2 y')^2)e^{x^2 y} + y'' = 0 \end{cases} .$$

Dalla prima di queste equazioni si ottiene $y'(x)$ in funzione di x e $y(x)$; sostituendo l'espressione ottenuta nella seconda equazione si ricava poi y''.
Ad esempio, per $x=0$, essendo $y(0)=-1$ si ottiene $y'(0)=1$, $y''(0)=2$.

Un altro metodo per ottenere le derivate di ordine superiore al primo è quello di derivare le derivate del primo ordine (assegnate da 7.6) pensate come funzioni di $\underline{x}$ composte mediante $y(\underline{x})$.
Ad esempio, per la funzione implicita definita da

$$x^4 + y^3 + x^2 + y^2 + 1 = 0$$

(vedasi par.5) si ha

$$y' = -\frac{4x^3+2x}{3y^2+2y}$$

e quindi

$$y'' = -\frac{(12x^2+2)(3y^2+2y) - y'(6y+2)(4x^3+2x)}{(3y^2+2y)^2} =$$
$$= -\frac{(12x^2+2)(3y^2+2y)^2 + (6y+2)(4x^3+2x)^2}{(3y^2+2y)^3} .$$

9. Teorema di esistenza e unicità in piccolo. Il caso m=1 .

Teorema 9.1 - La funzione $f: R^n \times R \to R$ sia definita e continua in un intorno I di un punto $(\underline{x}^\circ, y^\circ)$ tale che $f(\underline{x}^\circ, y^\circ)=0$; esista in tale intorno $\frac{\partial f}{\partial y}$ continua e sia inoltre $\frac{\partial f}{\partial y}(\underline{x}^\circ, y^\circ) \neq 0$. Allora è possibile determinare un intorno A di $\underline{x}^\circ$ in cui è definita una e una sola funzione $y(\underline{x})$ soddisfacente le condizioni seguenti

1) $y(\underline{x}^\circ)=y^\circ$

2) $(\underline{x}, y(\underline{x})) \in I \qquad \forall \underline{x} \in A$

3) $f(\underline{x}, y(\underline{x}))=0 \qquad \forall \underline{x} \in A$.

Inoltre, $y(\underline{x})$ è continua in A .

Tale funzione $y(\underline{x})$ viene chiamata "funzione definita implicitamente dall'equazione $f(\underline{x},y)=0$ a partire dal punto $(\underline{x}^\circ, y^\circ)$". Tale punto viene chiamato "punto origine" della funzione implicita.

Per dimostrare il teorema osserviamo che, in queste ipotesi, è possibile determinare un intervallo $A \subset R^n$ di centro $\underline{x}^\circ$ e un intervallo $B \subset R$ di centro y° tali che in $A \times B$ $\frac{\partial f}{\partial y}$ sia sempre positiva o sempre negativa. E' inoltre possibile (°) scegliere A e B abbastanza piccoli in modo che in $A \times B$ siano soddisfatte tutte le ipotesi del teorema 5.1. Da questo teorema e dal teorema 6.1 segue allora l'asserto.

Osservazioni - 1) L'esistenza di un punto origine $(\underline{x}^\circ, y^\circ)$ in cui f si annulla è indispensabile per la validità del teorema.

Ad esempio, la funzione $f(x,y)=x^2+e^y+1$ soddisfa tutte le altre ipotesi in $R \times R$, eppure l'equazione $f(x,y)=0$ non definisce implicitamente alcuna funzione $y(x)$.

2) Se nel punto origine $(\underline{x}^\circ, y^\circ)$ $\frac{\partial f}{\partial y}$ è nulla, non è possibile garantire nè l'esistenza di una funzione implicita nè, se essa esiste, la sua unicità.

Per esempio, l'equazione $x^2+y^2=0$ soddisfa in qualsiasi intorno di $(0,0)$ tutte le altre ipotesi del teorema, eppure non definisce implicitamente alcuna funzione $y(x)$ in un intorno di $x=0$; analogamente dicasi per l'equazione $y^2-x^2=0$ che invece definisce le quattro funzioni implicite continue $y(x)=x$, $y(x)=-x$, $y(x)=|x|$, $y(x)=-|x|$.

Dal teorema precedente e dal teorema 7.2 si ottiene immediatamente il

Teorema 9.2 - Valgano le ipotesi del Teorema 9.1 e

(°) Si veda la prima parte della dimostrazione del teorema 6.1.

inoltre f sia differenziabile in $(\underline{x}^\circ,y^\circ)$. Allora la funzione implicita $y(\underline{x})$ definita a partire da $(\underline{x}^\circ,y^\circ)$ è differenziabile in tale punto e

$$\operatorname{grad} y(\underline{x}^\circ) = \left(- \frac{\frac{\partial f}{\partial x_1}(\underline{x}^\circ,y^\circ)}{\frac{\partial f}{\partial y}(\underline{x}^\circ,y^\circ)}, \dots, - \frac{\frac{\partial f}{\partial x_n}(\underline{x}^\circ,y^\circ)}{\frac{\partial f}{\partial y}(\underline{x}^\circ,y^\circ)} \right).$$

Dal teorema 9.2 si ricava infine il

<u>Teorema</u> 9.3 - Valgano le ipotesi del Teorema 9.1 e inoltre $f \in \mathscr{C}^{(1)}(I)$. Allora la funzione implicita $y(\underline{x})$ definita a partire da $(\underline{x}^\circ,y^\circ)$ è di classe $\mathscr{C}^{(1)}(A)$ e

$$\operatorname{grad} y(\underline{x}) = \left(- \frac{\frac{\partial f}{\partial x_1}(\underline{x},y(\underline{x}))}{\frac{\partial f}{\partial y}(\underline{x},y(\underline{x}))}, \dots, - \frac{\frac{\partial f}{\partial x_n}(\underline{x},y(\underline{x}))}{\frac{\partial f}{\partial y}(\underline{x},y(\underline{x}))} \right) \qquad \forall\, \underline{x} \in A.$$

Ad esempio, l'equazione $x^3+y^3+x-y=0$ definisce implicitamente in un intorno di $x=0$ tre funzioni differenziabili, rispettivamente a partire dai punti $(0,0)$, $(0,1)$, $(0,-1)$.

10. <u>Applicazioni. Estremi liberi delle funzioni implicite</u>.

Considereremo, in questo paragrafo, funzioni $f: A\times B \to R$, con A aperto di R^n e B aperto di R e supporremo che f sia di classe $\mathscr{C}^{(k)}$ in $A\times B$, con k grande quanto sarà necessario.
I risultati dei paragrafi precedenti permettono di risolvere facilmente vari problemi relativi alle (eventuali) funzioni definite implicitamente dall'equazione $f(\underline{x},y)=0$.
Ad esempio, se il punto $(\underline{x}^\circ,y^\circ) \in A\times B$ è punto origine di una funzione implicita $y(\underline{x})$, tale funzione è di classe $\mathscr{C}^{(k)}$ in un intorno di $\underline{x}^\circ$ ed è possibile determinarne lo sviluppo di Taylor arrestato

al k-mo termine.

Consideriamo p.e. l'equazione $e^{x^2y}+y-x=0$. Abbiamo già visto (par.8) che per la funzione implicita $y=y(x)$ da essa definita a partire dal punto $(0,-1)$ si ha $y'(0)=1$, $y''(0)=2$. Di conseguenza, lo sviluppo di Taylor di questa funzione in un intorno dell'origine arrestato al termine di 2° ordine è

$$y(x) = -1+x+x^2+o(x^2) \quad .$$

Vediamo ora come sia possibile determinare i punti estremanti della funzione implicita. Vale il

<u>Teorema</u> 10.1 - Gli estremi liberi $(\underline{x},y(\underline{x}))$ delle funzioni differenziabili definite implicitamente dall'equazione $f(\underline{x},y)=0$ vanno ricercati fra le soluzioni del sistema

$$(10.1) \qquad \begin{cases} f(\underline{x},y)=0 \\ \frac{\partial f}{\partial x_i}(\underline{x},y)=0 \qquad i=1,\dots,n \end{cases} .$$

Infatti, sia $\underline{x}^\circ$ un punto estremante e $y(\underline{x}^\circ)=y^\circ$.
Allora $f(\underline{x}^\circ,y^\circ)=0$.
Inoltre, per il teorema 2.1, si ha anche $dy(\underline{x}^\circ)=0$.
Poichè da $f(\underline{x},y(\underline{x})) \equiv 0$ si ricava

$$\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\underline{x},y(\underline{x}))dx_i + \frac{\partial f}{\partial y}(\underline{x},y(\underline{x}))dy(\underline{x})=0$$

qualunque siano $dx_1,\dots,dx_n$, per $\underline{x}=\underline{x}^\circ$ si ottiene

$$\frac{\partial f}{\partial x_i}(\underline{x}^\circ,y(\underline{x}^\circ)) = 0 \qquad i=1,\dots,n$$

e vale quindi (10.1).

<u>Osservazioni</u> - 1) Il sistema (10.1) di n+1 equazioni nelle n+1 incognite $x_1,\dots,x_n,y$ permette di determinare le coordinate $x_1,\dots,x_n$ degli eventuali punti estremanti e, contemporaneamente, il corrispondente valore y della funzione implicita.

2) Se $n=1$, il sistema (10.1) si riduce al sistema di due equazioni in due incognite

$$\begin{cases} f(x,y)=0 \\ \dfrac{\partial f}{\partial x}(x,y)=0 \end{cases}$$

3) Sia $(\underline{x}^\circ, y^\circ)$ una soluzione del sistema (10.1). Ciò non assicura, in generale, l'esistenza di una funzione implicita differenziabile per la quale $y(\underline{x}^\circ)=y^\circ$.

Basta considerare p.e. la funzione $f(x,y)=x^2+y^2$ oppure la funzione $f(x,y) = x^2 - y^3$.

L'esistenza di una tale funzione implicita è però assicurata, per il teorema 9.1, qualora sia $\frac{\partial f}{\partial y}(\underline{x}^\circ, y^\circ)\neq 0$.

4) Sia $(\underline{x}^\circ, y^\circ)$ una soluzione del sistema (10.1), con $\frac{\partial f}{\partial y}(\underline{x}^\circ, y^\circ)=0$. In tale punto sono nulle tutte le derivate parziali della funzione f e quindi il teorema 9.1 non è applicabile nemmeno per ottenere dall'equazione $f(\underline{x},y)=f(x_1,\dots,x_n,y)=0$ la variabile x_i $(i=1,\dots,n)$ come funzione implicita delle altre e cioè di $x_1,\dots,x_{i-1},x_{i+1},\dots,x_n,y$. Un tale punto viene chiamato "punto singolare" per l'equazione $f(\underline{x},y)=0$.

Ad esempio, l'origine è punto singolare per le equazioni $x^2+y^2=0$, $x^2-y^2=0$, $x^5-y^2=0$. E' interessante osservare che, a partire da (0,0), la prima non definisce implicitamente alcuna funzione, la seconda definisce le due funzioni differenziabili $y=\mp x$, mentre la terza definisce due funzioni y(x) derivabili solo per $x \geq 0$ e definisce una e una sola funzione x(y) continua ma non differenziabile per $y=0$.

5) Sia $(\underline{x}^\circ, y^\circ)$ una soluzione del sistema (10.1) con $\frac{\partial f}{\partial y}(\underline{x}^\circ, y^\circ)\neq 0$. Per decidere se $\underline{x}^\circ$ è un punto di massimo o di minimo o un colle per la funzione implicita definita a partire da tale punto, si utilizzano i metodi indicati nel paragrafo 3 .

E s e m p i

1) Sia $f(x,y)=(x+y)^2-2x+y=0$. Gli eventuali estremi delle funzioni implicite definite da $f(x,y)=0$ sono assegnati dal sistema

$$\begin{cases} (x+y)^2-2x+y=0 \\ 2(x+y)-2=0 \end{cases} \quad \begin{cases} -2x+y=-1 \\ x+y=1 \end{cases} \quad \begin{cases} 3x=2 \\ y=1-x \end{cases} \quad \begin{cases} x=\frac{2}{3} \\ y=\frac{1}{3} \end{cases} .$$

Poichè $\frac{\partial f}{\partial y}=2(x+y)+1\neq 0$ in $(\frac{2}{3},\frac{1}{3})$, esiste effettivamente una funzione y(x) definita implicitamente a partire da $(\frac{2}{3},\frac{1}{3})$ e indefinitamente differenziabile in un intorno di $x=+\frac{2}{3}$, perchè $f\in\mathscr{C}^\infty(R\times R)$. Per tale funzione è $y(\frac{2}{3})=\frac{1}{3}$, $y'(\frac{2}{3})=0$ e inoltre

$$y''(x)=\frac{d}{dx}\left(-\frac{2(x+y-1)}{2x+2y+1}\right)=\frac{-6(1+y')}{(2x+2y+1)^2}$$

e quindi $y''(\frac{2}{3})=-\frac{2}{3}<0$. Di conseguenza $x=\frac{2}{3}$ è un punto di massimo.

2) Si consideri l'equazione

$$f(x_1,x_2,y)=e^{x_1x_2y}-x_1x_2^2y-2=0,\quad x_1,x_2,y\in R .$$

Gli eventuali estremi delle funzioni implicite $y=y(x_1,x_2)$ da essa definite sono da ricercarsi fra le soluzioni del sistema

$$\begin{cases} e^{x_1x_2y}-x_1x_2^2y-2=0 \\ x_2y\,e^{x_1x_2y}-x_2^2y=0 \\ x_1y\,e^{x_1x_2y}-2x_1x_2y=0 \end{cases}$$

che non ha alcuna soluzione. Quindi non esistono nè punti estremanti nè colli.

3) Si consideri l'equazione

$$f(x_1,x_2,y)=e^{x_1x_2y}-y=0 .$$

Gli eventuali estremi delle funzioni $y=y(x_1,x_2)$ in essa implicite vanno cercati fra le soluzioni del sistema

$$\begin{cases} e^{x_1x_2y}-y=0 \\ x_2y\,e^{x_1x_2y}=0 \\ x_1y\,e^{x_1x_2y}=0 \end{cases}$$

che ha l'unica soluzione (0,0,1) .

Per decidere se nell'origine $y(x_1,x_2)$ ha un estremo, occorre calcolare le derivate seconde in tale punto. Per ottenerle rapidamente, conviene procedere come segue.

Posto $\frac{\partial^2 y}{\partial x_1^2}(0,0)=a$, $\frac{\partial^2 y}{\partial x_1 \partial x_2}(0,0)=b$, $\frac{\partial^2 y}{\partial x_2}(0,0)=c$,essendo $y(x_1,x_2)$ di classe $\mathscr{C}^\infty$, dalla formula di Taylor si ottiene

$$y(x_1,x_2)=1+\frac{1}{2!}(ax_1^2+2bx_1x_2+cx_2^2)+o(\|\underline{x}\|^2) \quad .$$

Sostituendo nell'equazione data si ricava

$$e^{x_1x_2+o(\|\underline{x}\|^2)} - \frac{1}{2}ax_1^2 - bx_1x_2 - \frac{1}{2}cx_2^2 - 1 + o(\|\underline{x}\|^2) \equiv 0 \quad ,$$

da cui

$$1+x_1x_2-\frac{1}{2}ax_1^2-bx_1x_2-\frac{1}{2}cx_2^2 - 1 + o(\|\underline{x}\|^2) = 0$$

e quindi $a = c = 0$, $b = 1$. E' allora evidente che l'origine è un colle.

Terminiamo osservando che i risultati ottenuti permettono,in molti casi, di studiare, con le tecniche abituali,l'andamento di una funzione assegnata in forma implicita.

<u>E s e m p i</u>

1) Studiamo la funzione $y=y(x)$ implicita nell'equazione

$$x^3+y^3+x^2+y=0 \qquad x,y\in R \quad .$$

E' facile verificare che il teorema 5.1 è applicabile in $R\times R$ .Quindi l'equazione data definisce una e una sola funzione $y(x)$ per ogni $x\in R$, che è di classe $\mathscr{C}^\infty(R)$, perché f è indefinitamente differenziabile e $\frac{\partial f}{\partial y}\neq 0$.

Per $x\to\pm\infty$ dall'equazione data si ottiene $y^3\sim -x^3$ e cioè

$$y(x)\sim -x \quad \text{per} \quad x\to\mp\infty \quad ; \quad \lim_{x\to\pm\infty} y(x) = \mp\infty \quad .$$

Inoltre y si annulla per $x=0$ e $x=-1$.
Gli eventuali estremi sono assegnati dal sistema

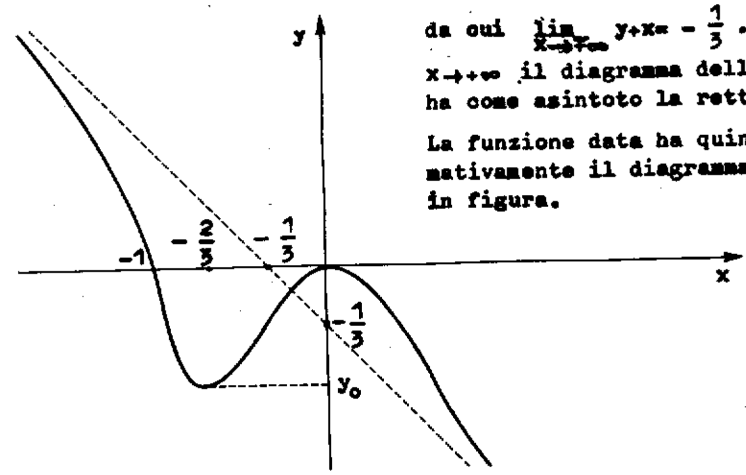
$$\begin{cases} 3x^2+2x=0 \\ x^3+y^3+x^2+y=0 \end{cases} \quad \text{da cui} \quad \begin{cases} x=0 \\ y=0 \end{cases} \quad e \quad \begin{cases} x=-\frac{2}{3} \\ y=y_* \end{cases}$$

dove y_* è la soluzione reale (unica!) dell'equazione $y^3+y+\frac{4}{27}=0$. E' evidente che $y_*<0$.

Tenuto conto del fatto che $y\to\pm\infty$ per $x\to\mp\infty$, $x=-\frac{2}{3}$ è punto di minimo e $x=0$ è punto di massimo.

Determiniamo ora gli eventuali asintoti. Abbiamo già visto che $y\sim -x$ per $x\to\mp\infty$. Inoltre dall'equazione data si ottiene anche

$$y+x=-\frac{x^2+y}{x^2-xy+y^2}$$



da cui $\lim_{x\to\mp\infty} y+x=-\frac{1}{3}$. Perciò per $x\to+\infty$ il diagramma della funzione y ha come asintoto la retta $y=-x-\frac{1}{3}$.

La funzione data ha quindi approssimativamente il diagramma indicato in figura.

2) Studiamo la funzione $y=y(x)$ implicita nell'equazione

$$f(x,y)=x\log y - y\,e^x = 0 \quad x\in R\ ,\ y\in R^+ \quad .$$

Consideriamo inanzi tutto $\frac{\partial f}{\partial y}=\frac{x}{y}-e^x$.

In $R^-\times R^+$, $\frac{\partial f}{\partial y}$ ha sempre segno negativo. Invece in $R^+\times R^+$ $\frac{\partial f}{\partial y}$ si annulla per $y=x\,e^{-x}$. Converrà quindi distinguere i due casi $x\in R^-$ e $x\in R^+$.

Per $x\in R^-$ risulta $\lim_{y\to 0+} f(x,y)=+\infty$, $\lim_{y\to+\infty} f(x,y)=-\infty$ e vale in $R^-\times R^+$ il teorema 5.1 .

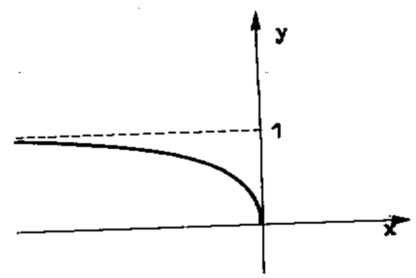
Per $x\in R^+$ risulta $\lim_{y\to 0+} f(x,y)=-\infty$, $\lim_{y\to+\infty} f(x,y)=-\infty$ ed inoltre $f(x,x\,e^{-x})=x(\log x-x-1)<0$. Quindi in $R^+\times R^+$ f non ha zeri.

Di conseguenza, l'equazione data definisce una e una sola funzione implicita $y(x)$ (di classe $\mathscr{C}^\infty$) per $x<0$ e nessuna per $x\geq 0$.

Poichè $x<0$ ed $y\,e^x>0$, affinchè l'equazione data sia soddisfatta deve essere $\log y<0$ e quindi

$$0<y(x)<1 \quad \forall\, x<0 \quad .$$

Poiché per $x\to-\infty$ $y\,e^x$ tende necessariamente a zero, deve essere anche $x\log y\to 0$ e quindi $\lim_{x\to-\infty} y(x)=1$; inoltre $\lim_{x\to 0-} y(x)=0$.



Poiché per $x\to-\infty$ $y\,e^x$ tende necessariamente a zero, deve essere anche $x\log y$ 0 e quindi $\lim_{x\to-\infty} y(x)=1$;

inoltre $\lim_{x\to 0-} y(x) = 0$.

Gli eventuali estremi di $y(x)$ sono assegnati dal sistema

$$\begin{cases} \log y - y e^{x} = 0 \\ x \log y - y e^{x} = 0 \end{cases}$$

che non ha soluzioni. Quindi $y(x)$ è sempre decrescente. Inoltre

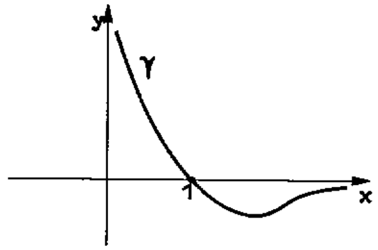
$$y'(x) = \frac{ye^{x} - \log y}{\frac{x}{y} - e^{x}} = \frac{ye^{x} - \log y}{e^{x}\left(\frac{1}{\log y} - 1\right)} \sim \log y \to -\infty \quad \text{per } x \to 0-$$

La funzione $y(x)$ ha quindi il diagramma approssimativo segnato in figura.

3) Studiamo l'andamento della funzione implicita definita, a partire dal punto $(0,1)$, dall'equazione

$$e^{xy} - x - y = 0 \quad .$$

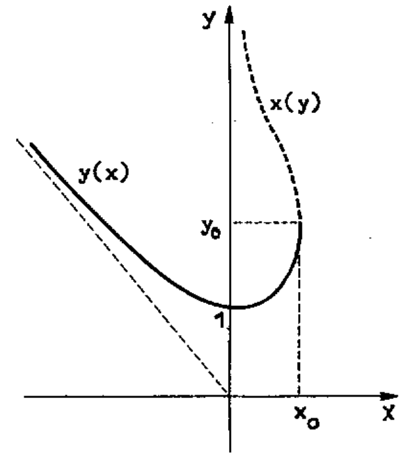
Poichè nel punto $(0,1)$ $\frac{\partial f}{\partial x} = y\, e^{xy} - 1 = 0$ e $\frac{\partial f}{\partial y} = x\, e^{xy} - 1 \neq 0$, tale punto è origine di una e una sola funzione implicita $y=y(x)$, di classe $\mathscr{C}^{\infty}$ in un opportuno intorno di $x=0$ e avente in tale punto derivata prima nulla. Cerchiamo di determinare il campo di esistenza di $x(y)$. I punti in cui $\frac{\partial f}{\partial y} = 0$ sono i punti del diagramma γ (indicato in figura) della funzione $y = -\frac{\log x}{x}$. Per quanto riguarda il semipiano $R^{-} \times R$ dal teorema 5.1 si ricava immediatamente che $y(x)$ è definita per ogni $x < 0$ ed è anzi in tale semipiano l'unica funzione y implicita nell'equazione data.



Inoltre $y \in \mathscr{C}^{\infty}(R^{-})$, è sempre positiva per $x<0$ e, per $x \to -\infty$, tende necessariamente a $+\infty$.
Osserviamo infine che il sistema

$$(*) \quad \begin{cases} \frac{\partial f}{\partial x} = y\, e^{xy} - 1 = 0 \\ f = e^{xy} - x - y = 0 \end{cases}$$

ha per $x \leq 0$ solo la soluzione $(0,1)$; quindi $y(x)$ è sempre decrescente nell'intervallo $(-\infty, 0)$.
Poichè $y''(0) > 0$, il punto $(0,1)$ è un minimo e y è crescente per $x>0$. (Si tenga presente che il sistema $(*)$ non ha soluzioni per $x>0$, $y>1$). Il suo diagramma deve perciò necessariamente incontrare la curva γ in un punto (x_o, y_o) con $0 < x_o < 1$. Poichè in tale punto la condizione $\frac{\partial f}{\partial x} = 0$ non è verificata, si ha $\lim_{x \to x_o^-} y'(x) = +\infty$; il diagramma di $y(x)$ si attesta in (x_o, y_o) con tangente verticale.
Poichè per il teorema 9.1 (x_o, y_o) è punto origine di una e una sola funzione implicita $x=x(y)$ che in tale punto ha un massimo, il campo di esistenza di $y(x)$ è $(-\infty, x_o)$ e il suo diagramma è quello indi-



cato in figura (dove è anche tratteggiata la parte rimanente del diagramma della funzione $x(y)$). Dalla figura si ricava anche che l'equazione data definisce, nell'intervallo $(0, x_o)$ due funzioni implicite continue di x.
Si osservi inoltre che la retta $y=-x$ è un asintoto della funzione $y(x)$ per $x \to -\infty$.
L'equazione data rimane inalterata scambiando fra loro x e y; quindi essa definisce implicitamente altre funzioni, il cui diagramma si ottiene da quello delle precedenti per simmetria rispetto alla retta $y=x$.

11. Teorema di esistenza e unicità in piccolo. Il caso $m > 1$.

Finora abbiamo studiato il problema delle funzioni implicite nel caso y reale. Il caso $\underline{y} \in R^{m}$ $(m>1)$ è notevolmente più complicato; tuttavia vale, anche in questo caso, un teorema di esistenza e unicità in piccolo analogo ai teoremi precedenti. Introduciamo, per semplicità di scrittura, le seguenti notazioni. Se $\underline{F} = \underline{F}(\underline{x}, \underline{y})$ è una funzione definita in $R^{n} \times R^{m}$ a valori in R^{p} $(n,m,p \geq 1)$, poniamo

$$P(\underline{x}, \underline{y}) = \left[\frac{\partial F_i}{\partial x_j}(\underline{x}, \underline{y})\right]_{\substack{i=1,\dots,p \\ j=1,\dots,n}} \qquad Q(\underline{x}, \underline{y}) = \left[\frac{\partial F_i}{\partial y_j}(\underline{x}, \underline{y})\right]_{\substack{i=1,\dots,p \\ j=1,\dots,m}}$$

Vale il

<u>Teorema</u> 11.1 - Siano $n,m \geq 1$, $\underline{F}: R^{n} \times R^{m} \to R^{m}$ una funzione di classe $\mathscr{C}^{(1)}$ in un intorno I di un punto $(\underline{x}^{\circ}, \underline{y}^{\circ}) \in R^{n} \times R^{m}$ in cui si annulla e sia inoltre $\det Q(\underline{x}^{\circ}, \underline{y}^{\circ}) \neq 0$. Allora è possibile determinare un intorno A di $\underline{x}^{\circ}$ in cui è definita una e una sola funzione $\underline{y}(\underline{x})$ a valori in R^{m} soddisfacente le condizioni seguenti:

1) $\underline{y}(\underline{x}^{\circ}) = \underline{y}^{\circ}$;
2) $(\underline{x}, \underline{y}(\underline{x})) \in I \quad \forall \underline{x} \in A$;

3) $\underline{F}(\underline{x},\underline{y}(\underline{x}))=\underline{0} \qquad \forall \underline{x}\in A$.

Inoltre tale funzione è di classe $\mathscr{C}^{(1)}(A)$ e

(11.1) $$J(\underline{y};\underline{x}) = -Q^{-1}(\underline{x},\underline{y}(\underline{x}))\cdot P(\underline{x},\underline{y}(\underline{x})) \quad .$$

Dimostreremo il teorema per induzione. Esso è vero per m=1 qualunque sia n. Sia $m>1$; proveremo che esso, se è vero per funzioni $\underline{F}$ a valori in $R^1 \ldots R^{m-1}$, è anche vero per funzioni $\underline{F}$ a valori in R^m.

Fra i minori di ordine m-1 di $P_{\underline{y}}(\underline{x}^\circ,\underline{y}^\circ)$ ne esiste necessariamente almeno uno a determinante non nullo: sia esso p.e. $J\binom{F_1\ldots F_{m-1}}{y_1\ldots y_{m-1}}$. Consideriamo i vettori ausiliari $\underline{u}\in R^{n+1}$ di componenti $(x_1,\ldots,x_n,y_m)$ e $\underline{v}\in R^{m-1}$ di componenti $y_1,\ldots,y_{m-1}$. Sia $\underline{\varphi}(\underline{u},\underline{v})$ la funzione vettoriale a valori in R^{m-1} di componenti $F_1,\ldots,F_{m-1}$.

Nel punto $(\underline{u}^\circ,\underline{v}^\circ)$ $(\underline{u}^\circ=x_1^\circ,\ldots,x_n^\circ,\ y_m^\circ\ ;\ \underline{v}^\circ=y_1^\circ,\ldots,y_{m-1}^\circ)$ sono soddisfatte le condizioni del teorema 8.1 , valido per funzioni a valori in R^{m-1}. Quindi $\underline{\varphi}(\underline{u},\underline{v}) = \underline{0}$ definisce implicitamente a partire da $(\underline{u}^\circ,\underline{v}^\circ)$ una e una sola funzione $\underline{v}=\underline{v}(\underline{u})$ e questa funzione è differenziabile con continuità in un intorno del punto $(\underline{u}^\circ,\underline{v}^\circ)$.

Sostituiamo in $F_m(x_1,\ldots,x_n,\ y_1,\ldots,y_m)$ le componenti della funzione $\underline{v}(\underline{u})$ al posto di $y_1,\ldots,y_{m-1}$. Otteniamo così una funzione $\Phi(\underline{x},y_m)$ a valori in R. Questa funzione è nulla in $(\underline{x}^\circ,y_m^\circ)$ ed è differenziabile con continuità in un intorno di tale punto, essendo composta di funzioni differenziabili.

Dimostriamo ora che è anche $\dfrac{\partial\Phi}{\partial y_m} \neq 0$ in $(\underline{x}^\circ,y_m^\circ)$. Infatti risulta :

(11.2) $$\frac{\partial\Phi}{\partial y_m} = \sum_{k=1}^{m-1}\frac{\partial F_m}{\partial y_k}\frac{\partial v_k}{\partial y_m} + \frac{\partial F_m}{\partial y_m}$$

ed è anche (essendo $\underline{\varphi}(\underline{u},\underline{v}(\underline{u}))$ identicamente nulla)

(11.3) $$\underline{0} = \sum_{k=1}^{m-1}\frac{\partial\underline{\varphi}}{\partial y_k}\frac{\partial v_k}{\partial y_m} + \frac{\partial\underline{\varphi}}{\partial y_m} \quad .$$

Le (11.2) e (11.3) formano un sistema lineare di m equazioni nelle m incognite $\dfrac{\partial v_1}{\partial y_m},\ldots,\dfrac{\partial v_{m-1}}{\partial y_m}$, $\dfrac{\partial\Phi}{\partial y_m}$.

Il determinante dei coefficienti è

$$\begin{vmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_{m-1}} & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_{m-1}} & -1 \end{vmatrix} = -J\binom{F_1\ldots F_{m-1}}{y_1\ldots y_{m-1}} \neq 0$$

in $(\underline{x}^\circ,\underline{y}^\circ)$. Tale sistema ha quindi una sola soluzione e, in particolare, si ha

$$\frac{\partial\Phi}{\partial y_m} = -\frac{J\binom{F_1\ldots F_m}{y_1\ldots y_m}}{J\binom{F_1\ldots F_{m-1}}{y_1\ldots y_{m-1}}} \neq 0 \qquad \text{in} \quad (\underline{x}^\circ,\underline{y}^\circ) \ .$$

Poichè il teorema 11.1 vale per m=1, l'equazione $\Phi(\underline{x},y_m)=0$ definisce implicitamente a partire da $(\underline{x}^\circ,y_m^\circ)$ una ed una sola funzione $y_m=y_m(\underline{x})$, differenziabile con continuità in un intorno di tale punto. Ponendo $y_k=y_k(\underline{x})=v_k(\underline{x},y_m(\underline{x}))$ $(k=1,\ldots,m-1)$ si viene a definire una funzione vettoriale $\underline{y}=\underline{y}(\underline{x})$ soddisfacente tutte le condizioni richieste. Questa funzione è inoltre differenziabile con continuità in un intorno di $(\underline{x}^\circ,\underline{y}^\circ)$ (tali essendo $y_m(\underline{x})$ e $\underline{v}(\underline{x},y_m)$) ed il suo differenziale soddisfa la relazione

$$d\underline{F} = P_{\underline{x}}(\underline{x},y(\underline{x}))d\underline{x} + P_{\underline{y}}(\underline{x},y(\underline{x}))d\underline{y} = \underline{0}$$

da cui la (11.1).

Osservazioni - 1) La (11.1) può venire ricordata facilmente; infatti basta differenziare la relazione $\underline{F}(\underline{x},\underline{y})=0$ e ricavare $d\underline{y}$ dall'equazione vettoriale $P(\underline{x},\underline{y})d\underline{x}+Q(\underline{x},\underline{y})d\underline{y}=\underline{0}$ così ottenuta.

2) Da (11.1) si ricava che se $\underline{F}$ è di classe $\mathscr{C}^{(k)}$ $(k \geq 1)$ in I, la funzione implicita è di classe $\mathscr{C}^{(k)}$ in A.

Ad esempio, consideriamo il sistema

$$\begin{cases} \cos x_1y_1 - x_2 + y_2 = 0 \\ \operatorname{sen} y_1y_2 - x_1 - x_3 = 0 \end{cases} \qquad \begin{array}{l} \underline{x}=(x_1,x_2,x_3)\in(-1,1)^3 \\ \underline{y}=(y_1,y_2)\in[-1,1]^2 \end{array} .$$

Le funzioni a primo membro sono di classe $\mathscr{C}^{\infty}$ nell'insieme di definizione; inoltre

$$P(\underline{x},\underline{y}) = \begin{bmatrix} -y_1 \operatorname{sen} x_1y_1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

$$Q(\underline{x},\underline{y}) = \begin{bmatrix} -x_1 \operatorname{sen} x_1y_1 & 1 \\ y_2 \cos y_1y_2 & y_1 \cos y_1y_2 \end{bmatrix} .$$

Il sistema dato è soddisfatto nel punto $\underline{x}°=\underline{0}$, $\underline{y}°=(0,-1)$; inoltre $\det Q(\underline{0},\underline{y}°)=1\neq 0$ e i primi membri del sistema sono di classe $\mathscr{C}^{\infty}$ nell'insieme di definizione. Quindi il sistema dato definisce implicitamente a partire da $(\underline{x}°, \underline{y}°)$ una e una sola funzione $\underline{y}=\underline{y}(\underline{x})$ di classe $\mathscr{C}^{\infty}$. Per tale funzione $\underline{y}=(y_1,y_2)$ si ha $y_1(\underline{0})=0$, $y_2(\underline{0})=-1$ e inoltre

$$J(\underline{y};\underline{0}) = -Q^{-1}(\underline{0},\underline{y}°)\cdot P(\underline{0},\underline{y}°)$$

$$= -\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

$$= -\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Quindi nel punto $\underline{x}°=\underline{0}$ si ha

$$\frac{\partial y_1}{\partial x_1} = -1 , \quad \frac{\partial y_1}{\partial x_2} = 0 , \quad \frac{\partial y_1}{\partial x_3} = -1 ; \quad \frac{\partial y_2}{\partial x_1} = 0 , \quad \frac{\partial y_2}{\partial x_2} = 1 , \quad \frac{\partial y_2}{\partial x_3} = 0 .$$

12. Funzione inversa.

Teorema 12.1 - Sia $n \geq 1$ e sia $\underline{f}:R^n \to R^n$ una funzione di classe $\mathscr{C}^{(1)}$ in un intorno A di un punto $\underline{x}° \in R^n$ e sia inoltre il determinante jacobiano

(12.1) $$\det J(\underline{f};\underline{x}°) \neq 0 .$$

Allora è possibile determinare un intorno $B \subset R^n$ del punto $\underline{y}°=\underline{f}(\underline{x}°)$ in cui è definita una e una sola funzione $\underline{x}=\underline{x}(\underline{y})$ a valori in A tale che risulti

$$\underline{f}(\underline{x}(\underline{y})) = \underline{y} .$$

Inoltre tale funzione è di classe $\mathscr{C}^{(1)}(B)$ e

(12.2) $$J(\underline{x},\underline{y}) = J^{-1}(\underline{f};\underline{x}(\underline{y})) \qquad \forall\, \underline{y} \in B .$$

Infatti, la funzione

$$F(\underline{y},\underline{x}) = \underline{y}-f(\underline{x})$$

soddisfa, nel punto $(\underline{y}°,\underline{x}°)$ tutte le ipotesi del teorema del Dini rispetto alla variabile $\underline{x}$ e quindi definisce implicitamente $\underline{x}=\underline{x}(\underline{y})$ in un intorno di $\underline{y}°$ con le proprietà richieste.

Osservazioni - 1) La funzione $\underline{x}=\underline{x}(\underline{y})$ viene chiamata "funzione inversa" della funzione $\underline{y}=\underline{f}(\underline{x})$ e denotata con $\underline{f}^{-1}$.
Il teorema 12.1 afferma in sostanza che, nelle ipotesi fatte, la funzione f è localmente invertibile.

2) Se $n=1$ la (12.2) diviene

$$\frac{df^{-1}}{dy}(y) = \frac{1}{\frac{df}{dx}(f^{-1}(y))}$$

e cioè la ben nota formula di derivazione della

funzione inversa di una funzione di una variabile.

E s e m p i

1) La funzione $f: R \to R$ definita da $f(x) = \text{sen}\, x$ è localmente invertibile in un intorno di ogni punto $x \neq \frac{\pi}{2} + k\pi$, dove cioè $f'(x) = \cos x \neq 0$.

2) La funzione $\underline{f}: R^2 \to R^2$ definita da

$$\underline{f}(x_1, x_2) = (e^{x_1} - x_2,\ x_1 + e^{x_2})$$

è invertibile in un intorno di qualsiasi punto di R^2; infatti si ha

$$\det J(\underline{f};\underline{x}) = \begin{vmatrix} e^{x_1} & -1 \\ 1 & e^{x_2} \end{vmatrix} = e^{x_1+x_2} + 1 \neq 0 \quad \forall \underline{x} \in R^2.$$

Si ha inoltre

$$J(\underline{f}^{-1};\ \underline{f}(\underline{x})) = \frac{\begin{bmatrix} e^{x_2} & 1 \\ -1 & e^{x_1} \end{bmatrix}}{e^{x_1+x_2} + 1} .$$

Se $\underline{f}$ soddisfa le ipotesi del teorema 12.1 relativamente al punto $\underline{x}°$, diremo, d'ora in poi, che $\underline{f}$ è regolare in $\underline{x}°$. Dal teorema 12.1 si ricavano allora, in particolare, i seguenti corolari.

Corollario 12.1 - La funzione $\underline{f}: R^n \to R^n$ è localmente invertibile in un intorno di ogni punto $\underline{x}°$ in cui essa è regolare e la funzione inversa è regolare nel punto $\underline{f}(\underline{x}°)$.

Corollario 12.2 - Se la funzione $\underline{f}: R^n \to R^n$ è regolare in un punto $\underline{x}°$, esistono un intorno di $\underline{x}°$ e un intorno di $\underline{f}(\underline{x}°)$ fra i quali $\underline{f}$ stabilisce un diffeomeomorfismo (°).

(°) cioè una biiezione continua e differenziabile.

13. Dipendenza e indipendenza funzionale.

Sia A un aperto di R^n e $f_1, \dots, f_m$ m funzioni reali di classe $\mathcal{C}^{(1)}(A)$. Consideriamo la funzione $\underline{f}: R^n \to R^m$ definita da $\underline{f}(\underline{x}) = (f_1(\underline{x}), \dots, f_m(\underline{x}))$ e sia $B = \underline{f}(A)$.
Diremo che $f_1, \dots, f_m$ sono (funzionalmente) dipendenti in A se esiste una (almeno!) funzione $F: R^m \to R$ differenziabile in B con $dF \neq 0$ e tale che

$$F(\underline{f}(\underline{x})) = 0 \qquad \forall \underline{x} \in A .$$

In caso contrario, diremo che $f_1, \dots, f_m$ sono (funzionalmente) indipendenti in A.

Per esempio, in R^2 sono dipendenti

$$f_1(x_1, x_2) = \text{sen}^2 x_1 - 2\cos^2 x_2 ,$$
$$f_2(x_1, x_2) = \cos^2 x_1 + 2\cos^2 x_2$$

e sono indipendenti

$$f_1(x_1, x_2) = \text{sen}\, x_1 x_2 , \quad f_2(x_1, x_2) = e^{x_1 + x_2} .$$

Teorema 13.1 - Se $f_1, \dots, f_m$ sono dipendenti in A, allora $J(\underline{f};\underline{x})$ ha caratteristica minore di m in ogni punto di A.

Infatti da $F(\underline{f}(\underline{x})) = 0 \quad \forall \underline{x} \in A$ si ricava $d\, F \circ \underline{f} = 0 \quad \forall \underline{x} \in A$ e cioè

$$(13.1) \quad 0 = \sum_{i=1}^{m} \frac{\partial F}{\partial y_i}(\underline{f}(\underline{x})) \cdot \frac{\partial f_i}{\partial x_j}(\underline{x}) \quad \forall \underline{x} \in I , \quad j = 1, \dots, n .$$

Essendo $dF \neq 0$ in B, il sistema lineare omogeneo (13.1) di n equazioni nelle m incognite $\frac{\partial F}{\partial y_i}$ ha una soluzione non banale. Di conseguenza, la caratteristica della matrice

$$\left[\frac{\partial f_i}{\partial x_j}(\underline{x})\right]_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = J(\underline{f};\underline{x})$$

deve essere minore di m, c.d.d.

Corollario 13.1 - Se $J(\underline{f};\underline{x})$ ha caratteristica m in A, allora $f_1, \dots, f_m$ sono indipendenti in A.

Il teorema 13.1 è invertibile. Vale infatti il seguente teorema, che ci limitiamo ad enunciare.

Teorema 13.2 - Se in A $J(\underline{f};\underline{x})$ ha caratteristica p, allora fra $f_1,\dots,f_m$ esistono p funzioni indipendenti in A e tutte le altre dipendono da queste.

Corollario 13.2 - Se $m>n$, $f_1,\dots,f_m$ sono dipendenti in A.

Corollario 13.3 - Sia $n=m$. Allora condizione necessaria e sufficiente affinchè $f_1,\dots,f_n$ siano dipendenti in A è che sia

$$\det J(\underline{f};\underline{x}) = 0 \qquad \forall\, \underline{x}\in A \ .$$

14. Estremi vincolati.

Nei paragrafi 2 e 3 ci siamo occupati della ricerca degli estremi di una funzione reale di più variabili indipendenti, supponendo che queste variabili fossero libere di assumere qualsiasi valore corrispondente ai punti di un aperto $A\subset R^n$. Nelle applicazioni si presenta però sovente anche il problema di determinare gli estremi di una funzione reale di più variabili, non più totalmente indipendenti fra di loro, ma obbligate a soddisfare certe condizioni espresse in generale da una o più equazioni scalari (il cui numero m non superi n-1). In altri termini, si pone il problema di determinare gli estremi di una funzione di più variabili in un insieme I che non è più un aperto di R^n, ma è costituito dai punti di una ipersuperficie (se m=1) o dai punti intersezione di due o più ipersuperfici (se $m>1$) di R^n.
Questo problema, nella maggior parte dei casi, può venire schematizzato nella forma seguente. Siano $f:R^n\to R$ e $\underline{F}:R^n\to R^m$ $(m<n)$ due funzioni definite in un aperto $A\subset R^n$. Ci si propone di

determinare gli estremi della funzione $f(\underline{x})$ qualora $\underline{x}$, anzichè descrivere liberamente l'insieme A, sia vincolato a muoversi in modo da soddisfare la condizione $\underline{F}(\underline{x})=\underline{0}$. Gli estremi di f in questa situazione vengono chiamati "estremi vincolati" o "condizionati".
Se $\underline{x}^\circ$ è un punto di massimo o minimo libero per f in A e se inoltre $\underline{F}(\underline{x}^\circ)=\underline{0}$, allora $\underline{x}^\circ$ è anche un punto di massimo o minimo vincolato per f; il viceversa non è, in generale, vero.
Ad esempio, per la funzione $f(x,y)=y+x^2$ l'origine non è un punto estremante, benchè sia un punto di minimo vincolato con la condizione $y=0$.

Supponiamo che f e $\underline{F}$ siano di classe $\mathscr{C}^{(1)}(A)$ e che la caratteristica di $J(\underline{F};\underline{x})$ sia m in ogni punto di A. Allora il problema della ricerca degli estremi vincolati è riconducibile, mediante il teorema del Dini, a quello della ricerca degli estremi liberi di una opportuna funzione reale di n-m variabili.
Infatti, sia $\underline{x}^\circ$ un punto estremante per f con la condizione $\underline{F}(\underline{x})=\underline{0}$ e supponiamo che il minore di ordine m formato con le colonne $i_1,\dots,i_m$ di $J(\underline{F};\underline{x}^\circ)$ abbia determinante diverso da zero. Poniamo $\underline{\eta}=(x_{i_1},\dots,x_{i_m})$ e sia $\underline{\xi}=(x_{j_1},\dots,x_{j_{n-m}})$ il vettore formato dalle rimanenti variabili. Allora $\underline{x}^\circ=(\underline{\xi}^\circ,\underline{\eta}^\circ)$ con $\underline{\xi}^\circ=(x^\circ_{j_1},\dots,x^\circ_{j_{n-m}})$ e $\underline{\eta}^\circ=(x^\circ_{i_1},\dots,x^\circ_{i_m})$ e $f(\underline{x})=f(\underline{\xi},\underline{\eta})$.
Per il teorema del Dini, l'equazione $\underline{F}(\underline{\xi},\underline{\eta})=\underline{0}$ definisce implicitamente una e una sola funzione $\underline{\eta}=\underline{\eta}(\underline{\xi})$ a partire dal punto $(\underline{\xi}^\circ,\underline{\eta}^\circ)$ e la funzione $g:R^{m-n}\to R$ definita da $g(\underline{\xi})=f(\underline{\xi},\underline{\eta}(\underline{\xi}))$ ha necessariamente un estremo nel punto $\underline{\xi}^\circ$. Quindi, la ricerca degli estremi di f col vincolo $\underline{F}(\underline{x})=\underline{0}$ è ricondotta alla ricerca degli estremi liberi della funzione $g(\underline{\xi})$.

Cerchiamo, ad esempio, gli estremi della funzione $f:R^2\to R$ definita da $f(x,y)=x+y$ con la condizione $x^2+y^2=1$, $y>0$.
L'equazione data definisce una e una sola funzione implicita $y=y(x)$ a valori in R^+, e cioè la funzione $y(x)=\sqrt{1-x^2}$ nell'intervallo $(-1,1)$.
Il problema è quindi ricondotto alla determinazione degli estremi liberi della funzione $g(x)=x+\sqrt{1-x^2}$. Poichè tale funzione ha un massimo nel punto $x = 1/\sqrt{2}$, la funzione f assegnata ha, nel punto

$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, un massimo vincolato con la condizione assegnata.

E' però opportuno cercare un metodo che consenta la ricerca degli estremi vincolati di f senza la preventiva determinazione della funzione $\underline{\eta}(\underline{\xi})$ definita implicitamente da $\underline{F}(\underline{\xi},\underline{\eta})=\underline{0}$. Vale il

<u>Teorema</u> 14.1 - Se f e $\underline{F}$ sono di classe $\mathscr{C}^{(1)}(A)$ e $J(\underline{F};\underline{x})$ ha caratteristica m in A, allora gli estremi di f col vincolo $\underline{F}(\underline{x})=\underline{0}$ vanno ricercati fra i punti stazionari in A della funzione

$$F_{\underline{\lambda}}(\underline{x}) = f(\underline{x}) + \underline{\lambda}\,\underline{F}(\underline{x}) \qquad \underline{\lambda} \in R^m$$

soddisfacenti l'ulteriore condizione $\underline{F}(\underline{x})=\underline{0}$.

Infatti, sia $\underline{x}°$ un punto di estremo vincolato e sia, come sopra, $\underline{\eta}=\underline{\eta}(\underline{\xi})$ la funzione definita implicitamente da $\underline{F}(\underline{\xi},\underline{\eta})=\underline{0}$ a partire da $\underline{x}°=(\underline{\xi}°,\underline{\eta}°)$. Poniamo, per semplicità di scrittura:

$$P_{\underline{\xi}} = \left[\frac{\partial f}{\partial x_{j_k}}(\underline{x}°)\right]_{k=1,\dots,n-m} \qquad P_{\underline{\eta}} = \left[\frac{\partial f}{\partial x_{i_k}}(\underline{x}°)\right]_{k=1,\dots,m}$$

$$Q_{\underline{\xi}} = \left[\frac{\partial F_h}{\partial x_{j_k}}(\underline{x}°)\right]_{\substack{h=1,\dots,m\\ k=1,\dots,n-m}} \qquad Q_{\underline{\eta}} = \left[\frac{\partial F_h}{\partial x_{i_k}}(\underline{x}°)\right]_{\substack{h=1,\dots,m\\ k=1,\dots,m}}$$

Per ogni $\Delta\underline{\xi}$ abbastanza piccolo e $\Delta\underline{\eta}=\underline{\eta}(\underline{\xi}°+\Delta\underline{\xi})-\underline{\eta}°$, da $\underline{F}(\underline{\xi},\underline{\eta}(\underline{\xi}))\equiv\underline{0}$ si ricava

$$Q_{\underline{\xi}}\Delta\underline{\xi} + Q_{\underline{\eta}}\Delta\underline{\eta} = \underline{0} \quad .$$

Inoltre, essendo $\underline{x}°$ un punto estremante per $f(\underline{\xi},\underline{\eta}(\underline{\xi}))$ si ha anche

$$P_{\underline{\xi}}\Delta\underline{\xi} + P_{\underline{\eta}}\Delta\underline{\eta} = 0 \quad .$$

Poichè il sistema di m+1 equazioni nelle m componenti di $\Delta\underline{\eta}$

$$\begin{cases} Q_{\underline{\xi}}\Delta\underline{\xi}+Q_{\underline{\eta}}\Delta\underline{\eta}=\underline{0} \\ P_{\underline{\xi}}\Delta\underline{\xi}+P_{\underline{\eta}}\Delta\underline{\eta}=0 \end{cases}$$

ha soluzione, la matrice dei coefficienti deve avere caratteristica al più m. Essendo $\det Q_{\underline{\eta}}\neq 0$, deve quindi esistere un m-vettore $\underline{\lambda}$ tale che

$$P_{\underline{\xi}} = \underline{\lambda}\, Q_{\underline{\xi}} \quad , \quad P_{\underline{\eta}} = \underline{\lambda}\, Q_{\underline{\eta}}$$

e cioè

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^{m} \lambda_j \frac{\partial F_j}{\partial x_i} \quad .$$

Di conseguenza, la funzione

$$F_{\underline{\lambda}} = f - \underline{\lambda}\,\underline{F}$$

ha differenziale nullo nel punto $\underline{x}°$, c.d.d.

<u>Osservazione</u> - Nelle ipotesi fatte, gli estremi vincolati di f vanno ricercati fra le soluzioni del sistema di n+m equazioni nelle n+m incognite $x_1,\dots,x_n$, $\lambda_1,\dots,\lambda_m$:

$$\begin{cases} d(f+\underline{\lambda}\,\underline{F})(\underline{x})=0 \\ \underline{F}(\underline{x})=\underline{0} \end{cases}$$

e cioé

$$(14.1)\quad \begin{cases} \dfrac{\partial f}{\partial x_i}(\underline{x}) + \sum_{j=1}^{m}\lambda_j \dfrac{\partial F_j}{\partial x_i}(\underline{x})=0 & i=1,\dots,n \\ F_j(\underline{x}) = 0 & j=1,\dots,m \end{cases} .$$

Il vettore $\underline{\lambda}$ viene chiamato "moltiplicatore di Lagrange".

<u>Esempi</u>

1) Determiniamo i punti estremanti della funzione $e^{x^2-y^2}$ sulla circonferenza $x^2+y^2=1$ $(x,y\in R)$.
Basta evidentemente determinare gli estremi della funzione ausiliaria $f(x,y)=x^2-y^2$.
La matrice jacobiana della funzione $F(x,y)=x^2+y^2-1$ ha caratteristica 1 in ogni punto della circonferenza. Per il teorema 14.1 i punti cercati sono allora fra le soluzioni del sistema (14.1), cioé del sistema

$$\begin{cases} 2x(1+\lambda)=0 \\ 2y(1-\lambda)=0 \\ x^2+y^2-1=0 \end{cases} .$$

Questo sistema ha le soluzioni $x=0$, $y=\mp 1$, $\lambda=1$ e $x=\mp 1$, $y=0$, $\lambda=-1$. Poichè f deve avere necessariamente almeno un minimo e un massimo sulla circonferenza data (è continua su un compatto!) e per motivi di simmetria, i punti $(0,\mp 1)$ sono minimi e i punti $(\mp 1,0)$ sono

massimi per f (e quindi anche per $e^{x^2-y^2}$) con il vincolo assegnato.

2) Determiniamo i punti estremanti della funzione $f(x,y,z)=x-z$ sulla circonferenza γ di R^3 di equazioni

$$\begin{cases} x^2+y^2+z^2=1 \\ x+y+z=0 \end{cases} .$$

In questo caso $\underline{F}: R^3 \to R^2$ ha come componenti $F_1(x,y,z)=x^2+y^2+z^2-1$ e $F_2(x,y,z)=x+y+z$. La sua matrice jacobiana

$$\begin{bmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{bmatrix}$$

ha caratteristica minore di 2 solo sulla retta $x=y=z$, che non incontra γ. Il sistema (14.1) in questo caso diviene

$$\begin{cases} 1+2\lambda x+\mu=0 \\ 2\lambda y+\mu=0 \\ -1+2\lambda z+\mu=0 \\ x^2+y^2+z^2=1 \\ x+y+z=0 \end{cases}$$

che ha le soluzioni $x=-\frac{1}{\sqrt{2}}$, $y=0$, $z=\frac{1}{\sqrt{2}}$, $\lambda=\frac{1}{\sqrt{2}}$, $\mu=0$ e $x=\frac{1}{\sqrt{2}}$, $y=0$, $z=-\frac{1}{\sqrt{2}}$, $\lambda=-\frac{1}{\sqrt{2}}$, $\mu=0$.

Essendo f continua e γ compatto, il punto $(-\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}})$ è un minimo e $(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}})$ un massimo per f su γ.